Zoological Research

ZOOLOGICAL RESEARCH

Volume 41, Issue 1   18 January 2020

CONTENTS

Cover image: *Macaca fascicularis*. Photo by Xiao-Feng Ma
Cover design: Li-Bin Wu

Nothing but the facts: a journey together

Peer review plays an essential role in academic quality control of scientific journals and is central to scholarly communication. It exists in many aspects of academia, including funding applications, career building, academic performance evaluation, and publishing. In regard to publishing, different forms of publication, from traditional hardcopies and on-line journals to preprint articles (e.g., arXiv: https://arxiv.org/, BioRxiv: https://www.biorxiv.org/), continue to fine-tune the requirements of peer review, and vice versa, peer review continuously refines the quality of publications.

In 2015, ORCID (https://orcid.org/), ScienceOpen (https://www.scienceopen.com/), Sense About Science (https://senseaboutscience.org/), and Wiley (https://www.wiley.com/en-us) initiated the annual global event, "Peer Review Week" (Meadows, 2019). By 2019, Peer Review Week has covered a wide variety of topics, including "Recognition for Review", "Transparency in Peer Review", "Diversity in Peer Review", and "Quality in Peer review". These are the same issues often discussed and considered by the Editorial Board of *Zoological Research* (*ZR*).

ZR adopts a journal-facilitated traditional single-blind pre-publication peer-review process, in which editors mediate all interactions between reviewers and authors; peer reviews are not published, and reviews are owned by the authors of said reviews. At the beginning of a new year, *ZR* would like to sincerely thank all reviewers for their rigorous work and dedication, particularly for maintaining the high standard of articles and helping *ZR* remain a useful literature resource. *ZR* recognizes our reviewers' contributions through the public acknowledgement column on the homepage of *ZR* (http://www.zoores.ac.cn/EN/column/item32.shtml), which is updated bi-monthly, and via the "annual personnel contribution award".

Here, we would like to invite all readers and authors of *ZR* to peruse some of our peer-review statistics. Before 2014, the year in which *ZR* began to publish in English only, we maintained a peer-review expert group of approximately 2 000 people, primarily Chinese scholars. After 2014, especially

since the introduction of our online submission system (ScholarOne Manuscripts, https://mc03.manuscriptcentral.com/zr) in September 2015, the total number of peer reviewers for *ZR* has doubled and now covers 46 countries worldwide.

From September 2015 to December 2019, *ZR* has received 610+ submissions, with 206 having been published in 25 issues (Vol.36(6)–Vol.40(6)). This has resulted in an overall rejection rate of approximately 67%. After the initial change in publishing language, which impacted our authorship and readership, *ZR* attempted to solicit new submissions, even running ScholarOne Manuscripts (for English submissions) and JournalX of Beijing Magtech S&T Co., Ltd (for Chinese submissions) parallelly for over 18 months. Despite this, the independent submission quantity dropped in 2015 and 2016. Thus, 61% (26 out of 43) and 74% (31 out of 42) of publications in 2017 and 2018 were invited contributions, respectively. Thanks to the enduring faith and generous support of readers and authors, *ZR* has now experienced a submission upturn, with overall submissions in 2019 the sum of those in 2017 and 2018 combined.

At present, more than 55% of submissions are rejected during the preliminary review stage conducted by a subject-expert editorial board member and/or journal editor. Excluding this proportion, or the ones are still under peer reviewing, a total of 200 manuscripts have gone through peer review via ScholarOne Manuscripts. Regarding these papers, invitations for review were sent to 1 951 scholars, with 607 accepting the invitation and completing the reviewing process. The average time and review report length for these submissions was 19 days and 476 words, respectively. Among the 607 reviewers, 33.7% (205) were from China, 11.9% (72) were from the US, followed by Japan (26, 4.3%), Australia (13, 2.1%), France, UK, and India (11, 1.8%).

Based on the survey of Vesper (2018), *ZR* does experience "reviewer fatigue" during the invitation process. On average, for each review to be completed, we must invite 3.2 scholars. However, when using review time and average report length as two parameters of efficiency, we currently achieve the global median level (19 vs. 16.4 days and 476 vs. 447 words, respectively) (Global State of Peer Review, 2018).

Thus, we are happy to say that, with the support of our readers, authors, and reviewers, *ZR* is on the right track in regard to our submissions and review process. Our articles cover classical aspects to cutting edge research within the zoological field. We thank the authors for trusting *ZR* with their

discoveries and the reviewers for generously sharing their expertise with us.

To maintain the objectivity, independency, and efficiency of peer review, *ZR* endeavors to acquire feedback from the broader biological community by balancing reviewer gender and by selecting reviewers based on diverse geographic areas and career stages. To help in this, we kindly encourage invited reviewers and authors to consider the above objectives when recommending alternative reviewers or scholars during the submission process. *ZR* follows the Ethical Guidelines for Peer Reviewers of Committee on Publication Ethics (COPE Council, 2017). If you wish to become more involved in the peer review of *ZR* or have any suggestions on what we can do better, please feel free to contact us at zoores@mail.kiz.ac.cn.

Thanks to the growing impact of *ZR* (Yao et al., 2019), we have seen a significant increase in submissions during the past year. Many publications, such as the Chinese tree shrew genome analysis (Fan et al., 2019), study of ecology and social system of northern gibbons (Guan et al., 2018), roles of ayu NOD2 in antibacterial innate immunity (Ren et al., 2019), pattern of allele-specific expression and alternative splicing in hybrids of domestic animals (Wang et al., 2019), and others, have received considerable attention in the field. 2020 is the 40th anniversary of the founding of *ZR*. We believe that with the full support of our authors, reviewers, editorial board members, and readers, *ZR* will continue its collaborative journey and achieve the status of a top journal in Zoology.

Suqing Liu

Su-Qing Liu, Editorial Director of *Zoological Research*

Kunming Institute of Zoology, Chinese Academy of Sciences
Kunming, Yunnan 650223, *China*

REFERENCES

COPE Council. 2017. Ethical guidelines for peer reviewers. www.publicationethics.org.

Fan Y, Ye MS, Zhang JY, Xu L, Yu DD, Gu TL, Yao YL, Chen JQ, Lv LB, Zheng P, Wu DD, Zhang GJ, Yao YG. 2019. Chromosomal level assembly and population sequencing of the Chinese tree shrew genome. *Zoological Research*, **40**(6): 506−521.

Global State of Peer Review, 2018. https://publons.com/community/gspr; https://publicationethics.org/resources/guidelines-new/cope-ethicalguidelines-peer-reviewers.

Guan ZH, Ma CY, Fei HL, Huang B, Ning WH, Ni QY, Jiang XL, Fan PF. 2018. Ecology and social system of northern gibbons living in cold seasonal forests. *Zoological Research*, **39**(4): 255−265.

Meadows A. 2019. Peer review week is five! https://scholarlykitchen.sspnet.org/2019/06/26/peer-review-week-is-five/.

Ren Y, Liu SF, Nie L, Cai SY, Chen J. 2019. Involvement of ayu NOD2 in NF-κB and MAPK signaling pathways: Insights into functional conservation of NOD2 in antibacterial innate immunity. *Zoological Research*, **40**(2): 77−88.

Vesper I. 2018. Peer reviewers unmasked: largest global survey reveals trends. https://www.nature.com/articles/d41586-018-06602-y.

Wang Y, Gao S, Zhao Y, WH Chen, Shao JJ, NN Wang, Li M, Zhou GX, Wang L, Shen WJ, Xu JT, Deng WD, Wang W, Chen YL, Jiang Y. 2019. Allele-specific expression and alternative splicing in horse×donkey and cattle×yak hybrids. *Zoological Research*, **40**(4): 293−305.

Yao YG, Zhang Y, Zheng YT. 2019. An "impact" in publishing. *Zoological Research*, **40**(4): 239−240.

Neuroprotectants attenuate hypobaric hypoxia-induced brain injuries in cynomolgus monkeys

Pei Zhang[1,2,3,#], Jie-Si Chen[4,#], Qi-Ye Li[2,#], Long-Xiang Sheng[4,#], Yi-Xing Gao[5,6,7,#], Bing-Zheng Lu[4], Wen-Bo Zhu[4], Xiao-Yu Zhan[2], Yuan Li[4], Zhi-Bing Yuan[5,6,7], Gang Xu[5,6,7], Bi-Tao Qiu[3], Min Yan[4], Chun-Xue Guo[2], You-Qiong Wang[4], Yi-Jun Huang[4], Jing-Xia Zhang[8], Fu-Yu Liu[5,6,7], Zhong-Wei Tang[5,6,7], Sui-Zhen Lin[9], David N. Cooper[10], Huan-Ming Yang[2,11], Jian Wang[2,11], Yu-Qi Gao[5,6,7,*], Wei Yin[12,*], Guo-Jie Zhang[1,3,13,14,*], Guang-Mei Yan[4,*]

[1] *State Key Laboratory of Genetic Resources and Evolution*, *Kunming Institute of Zoology, Chinese Academy of Sciences*, *Kunming, Yunnan 650223, China*

[2] *BGI-Shenzhen, Shenzhen, Guangdong 518083, China*

[3] *Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen DK-2100, Denmark*

[4] *Department of Pharmacology, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, Guangdong 510080, China*

[5] *Institute of Medicine and Hygienic Equipment for High Altitude Region, College of High Altitude Military Medicine, Third Military Medical University, Chongqing 400038, China*

[6] *Key Laboratory of High Altitude Medicine of People's Liberation Army, Chongqing 400038, China*

[7] *Key Laboratory of High Altitude Environmental Medicine, Third Military Medical University, Ministry of Education, Chongqing 400038, China*

[8] *School of Pharmaceutical Sciences, Sun Yat-Sen University, Guangzhou, Guangdong 510006, China*

[9] *Guangzhou Cellprotek Pharmaceutical Co. Ltd., Guangzhou, Guangdong 510663, China*

[10] *Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK*

[11] *James D. Watson Institute of Genome Sciences, Hangzhou, Zhejiang 310058, China*

[12] *Department of Biochemistry, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou, Guangdong 510080, China*

[13] *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

[14] *China National Genebank, BGI-Shenzhen, Shenzhen, Guangdong 518120, China*

ABSTRACT

Hypobaric hypoxia (HH) exposure can cause serious brain injury as well as life-threatening cerebral edema in severe cases. Previous studies on the mechanisms of HH-induced brain injury have been conducted primarily using non-primate animal models that are genetically distant to humans, thus hindering the development of disease treatment. Here, we report that cynomolgus monkeys (*Macaca*

fascicularis) exposed to acute HH developed human-like HH syndrome involving severe brain injury and abnormal behavior. Transcriptome profiling of white blood cells and brain tissue from monkeys exposed to increasing altitude revealed the central role of the HIF-1 and other novel signaling pathways, such as the vitamin D receptor (VDR) signaling pathway, in co-regulating HH-induced inflammation processes. We also observed profound transcriptomic alterations in brains after exposure to acute HH, including the activation of angiogenesis and impairment of aerobic respiration and protein folding processes, which likely underlie the pathological effects of HH-induced brain injury. Administration of progesterone (PROG) and steroid neuroprotectant 5α-androst-3β,5,6β-triol (TRIOL) significantly attenuated brain injuries and rescued the transcriptomic changes induced by acute HH. Functional investigation of the affected genes suggested that these two neuroprotectants protect the brain by targeting different pathways, with PROG enhancing erythropoiesis and TRIOL suppressing glutamate-induced excitotoxicity. Thus, this study advances our understanding of the pathology induced by acute HH and provides potential compounds for the development of neuroprotectant drugs for therapeutic treatment.

INTRODUCTION

High altitude, which represents one of the most extreme environments on Earth, creates hypobaric hypoxia (HH) conditions to which only a small proportion of people can adapt. As the brain is the most hypoxia-intolerant organ, most people who ascend too rapidly to high altitudes cannot acclimatize to the HH environment and frequently suffer from a range of symptoms caused by acute hypoxia (Wilson et al., 2009). In severe cases, patients can develop serious high-altitude cerebral edema (HACE). With over 35 million people around the world traveling to high-altitude regions annually for tourism, sport, or work (Martin & Windsor, 2008), brain damage caused by HH has become a health hazard for lowland people who move to high altitudes. So far, the main mechanisms underlying brain damage include metabolic disturbance of neural cells, increased permeability of brain microvasculature, and oxidative stress (Basnyat & Murdoch, 2003; Wilson et al., 2009). However, details on the cascades and gene regulatory networks through which brain damage is induced by HH have not yet been determined.

Several drugs, such as acetazolamide and dexamethasone, are proposed to treat or even prevent HH-induced brain damage and edema based on increased oxygen-carrying capacity or anti-inflammatory effects (Imray et al., 2010). However, their efficacy is still controversial and numerous side effects have been reported in relation to their clinical use (Imray et al., 2010; Rabinstein, 2006). Progesterone (PROG), a well-known endogenous neuroprotective sex hormone, is reported to be associated with acute HH acclimatization (Stein, 2008). In addition, serum levels of PROG are found to increase in men who travel to high-altitude areas (Basu et al., 1997), suggesting the possibility that PROG may ameliorate HH-induced brain damage. Furthermore, as first reported by our group (Chen et al., 2013a), 5α-androst-3β,5,6β-triol (TRIOL) is a novel neuroprotectant and analogue of the endogenous neuroprotective steroid cholestane-3β,5α,6β-triol (Hu et al., 2014), thus its effects on HH-induced brain damage are also worth investigation.

Previous pathological and molecular studies of HH-induced brain injuries have focused primarily on small mammals such as mice and rats (Imray et al., 2010). However, these rodents display distinct hypoxia tolerance and physiological responses to HH in comparison to humans. For example, both acute HH conditions and exhaustive exercise are necessary to induce HACE-like symptoms in rats (Guo et al., 2013), and very rapid ascent speed (velocity of 50 m/s within 5 min to 6 000 m) is required to induce HACE-like symptoms in mice (Huang et al., 2015). In addition, the potential drug targets in rodents may differ from those in humans due to the large genetic differences between rodents and primates after ~96 million years of evolutionary divergence (Nei et al., 2001). In particular, the central nervous system (CNS) of primates is much more complex than that of rodents (Lissa et al., 2013). Thus, given the limitations of rodent models, the establishment of non-human primate models for HH study is urgently needed and should provide valuable insights into the molecular cascades and gene regulatory networks underlying HH-induced brain diseases and further our understanding of the effects and molecular mechanisms of neuroprotectants as potential drugs.

In the current study, we elucidated the spatial and temporal influence of acute HH on gene expression and examined the possible effects of PROG and TRIOL on HH-induced brain damage *in vivo*. Specifically, we established a non-human primate model of HH-induced brain damage and profiled the transcriptomes of white blood cells (WBCs) and brain tissue from cynomolgus monkeys before and after acute HH exposure (i.e., increasing altitude). Extensive gene regulatory analyses revealed a dynamic change in the WBC transcriptome response to HH as well as the gene regulatory network of newly identified transcriptomic hub genes. Moreover, the application of the two neuroprotectants effectively protected the brains of cynomolgus monkeys from HH-induced injury. Finally, PROG and TRIOL exerted their effects via different pathways, the former through erythropoiesis and the latter by suppression of glutamate-induced excitotoxity.

MATERIALS AND METHODS

Experimental animals

All experiments were conducted in accordance with the Chinese Laws for the Protection of Animals. The experimental protocols were approved by the Ethics Committee of Zhongshan School of Medicine, Sun Yat-Sen University according to the ARRIVE (Animal Research: Reporting of *In Vivo* Experiments) guidelines (Kilkenny et al., 2010). All animal-based procedures were performed in strict adherence to the National Standards of Treating Experimental Animals (2006 version). Efforts were taken to minimize suffering and to ensure the welfare of monkeys during experimentation.

Twenty-four male cynomolgus monkeys (6.0–6.5 years old, 6.8–7.5 kg) were obtained from Gaoyao Kangda Laboratory Animals Science & Technology Co., Ltd, Zhaoqing, Guangdong Province, China. The animals were transported to the Third Military Medical University, Chongqing, China. Monkeys were singly housed in cages in a controlled environment at a temperature of 25±1 °C, relative humidity of 60%, and circadian 12 h light/dark cycle and were fed routinely.

For experimentation, monkeys were randomly divided into four groups, each containing six individuals. The first group was maintained under normobaric normoxia (NN) conditions (at an altitude of 320 m) as a control. The other three groups were subjected to hypobaric hypoxia (HH), as detailed below. Groups 3 and 4 were treated with progesterone (HH+PROG) and 5α-androst-3β,5α,6β-triol (HH+TRIOL), respectively.

After 48 h at a simulated altitude of 7 500 m, the HH monkeys were sacrificed. The NN monkeys were also sacrificed after 3 d in the NN environment. All monkeys were euthanized by bloodletting from the carotid under anesthesia using a mixture of injectable ketamine hydrochloride (0.06 mg/kg) and xylazine hydrochloride (0.02 mg/kg).

Pharmacological treatments

Intravenous (10 mg/mL) and extended-release intramuscular injections (50 mg/mL) of TRIOL were provided by Guangzhou Cellprotek Pharmaceutical Company Co., Ltd. (China). The TRIOL (Batch No. 101124) was dissolved in vehicle containing 0.9% sodium chloride and 20% hydroxypropyl-β-cyclodextrin (HP-β-CD) for intravenous injection, and in vehicle containing 12% glycerin, 20% HP-β-CD, and 0.19% CMC-Na for extended-release intramuscular injection. Injectable progesterone (20 mg/mL) was purchased from Zhejiang Xianju Pharmaceutical Co., Ltd. (China), with a CFDA ratification No. of Guo Yao Zhun Zi-H33020828. The HP-β-CD (Batch No. 100309) was purchased from Xi'an Deli Biology & Chemical Industry Co., Ltd. (China). Injectable ketamine hydrochloride (150 mg/mL) was purchased from Shenyang Veterinary Drugs Co., Ltd. (China) with an approval No. of (2011)060022668 by the Ministry of Agriculture of the People's Republic of China (PRC). Injectable xylazine hydrochloride (100 mg/mL) was purchased from Dunhua Shengda Veterinary Drugs Co., Ltd. (China) with an approval No. of (2010)070031581 by the Ministry of Agriculture of the PRC.

Acute hypobaric hypoxia and neuroprotective steroid agent treatments

For HH treatment, the simulated altitude in the hypobaric chamber was increased from 320 m to 6 000 m at a velocity of 3 m/s, and from 6 000 m to 7 500 m at a velocity of 2 m/s. Altitude simulation was started at 320 m (stage A), then suspended at 3 000 m (stage B), 4 500 m (stage C), and 6 000 m (stage D) for 50 min each, and finally maintained at 7 500 m for an initial 24 h (stage E) and then another 24 h (stage F) for a total of 48 h. At every stage, monkeys were maintained under HH conditions for 30 min at the corresponding altitudes, and were then subjected to blood sample collection and drug administration within 20 min. Because of safety concerns for the experimenters, the altitude was decreased from 7 500 m to 6 000 m for blood collection and drug administration at stage E within 30 min. Throughout the HH experimental exposure, the temperature, relative humidity, and air-flow velocity of the chamber were maintained at 22 °C, 60%, and 5 L/min, respectively.

For the HH+PROG group, animals were intramuscularly injected with 15 mg/kg of PROG three times, first at 12 h before HH treatment, then during stages D and E. For the HH+ TRIOL group, animals were intravenously injected with 10 mg/kg of TRIOL 5 min before HH treatment, and at stages B, C, D, and E. Owing to TRIOL's short half-life ($t_{1/2}$) of 0.5 h, the extended-release intramuscular injection of TRIOL (30 mg/kg) was also administered at stages D and E.

Assessment of skeletal muscle coordination behavior

Behavioral assessment was performed by trained experimenters blind to group information. Monkey behavior was monitored 2 h after arrival at the simulated altitude of 7 500 m. We observed their activities over the next 30 min and employed a modified scale method to assess skeletal muscle coordination, with higher scores indicating more severely impaired behavioral states. The scoring system for monkey behavior assessment is presented in Supplementary Table S1.

Measurement of brain water content

Left brain hemispheres were used to measure brain water content with a precision electronic scale (Sartorius, BSA224S, Germany). Briefly, whole left hemispheres were immediately removed and measured as 'wet weight'. The left hemispheres were then placed in an electric thermostatic oven at 60 °C and weighed repeatedly every day until reaching a constant value (dry weight). The percentage of brain water content was calculated as: water content (%)=(wet weight-dry weight)/wet weight×100%.

Transmission electron microscopy (TEM)

The frontal cortices of brains were divided into 1 mm×1 mm× 1 mm pieces and fixed in 2.5% glutaraldehyde overnight at 4 °C. The fixed brain tissues were dehydrated through an ethanol series, embedded in epoxy resin, and cut into 60 nm ultrathin sections with a Leica EM UC6 Ultramicrocut (Germany). The sections were mounted on copper grids and

then stained in uranyl acetate and citric acid. Images were captured using a transmission electron microscope (Tecnai G2 Spirit Twin, FEI, USA). The experimental procedures for TEM were completed in a double-blind manner.

Nissl staining

The frontal cortices of brains were cut into 5 mm-thick pieces immediately after removal and then fixed in 4% paraformaldehyde. All samples were dehydrated in an increasing series of ethanol, cleared in xylene, and embedded in paraffin. Brain sections of 5 μm-thickness were cut using a rotary microtome (Shandon Finesse 325, Thermofisher Scientific, USA) and then deparaffinized with xylene and hydrated, followed by Nissl staining (0.5% cresyl violet) for histopathological assessment of neuronal injury. For Nissl staining, injured neurons were characterized by dark staining, condensed nuclei, shrunken cell bodies, or weak staining with irregular shapes. Images were acquired with a bright-field microscope (Nikon ECLIPSE Ti-U, Japan).

WBC isolation

WBCs were isolated from whole blood samples using Red Cell Lysis Buffer (Tiangen, RT122, China). In brief, red blood cells (RBCs) were lysed by gently mixing the blood sample with Red Cell Lysis Buffer (1:3), followed by incubation on ice for 5 min. The precipitate was then harvested by centrifugation at 2 500 g for 5 min at 4 °C. The WBC samples were ready for total RNA extraction after the isolation procedure was repeated three times.

RNA extraction, library construction, and sequencing

Total RNA was extracted and purified from WBCs and frontal cortices using Trizol reagent (Invitrogen, 15596, USA) according to the manufacturer's standard protocols. RNA quality and concentration were assessed using an Agilent Bioanalyzer 2100 and RNA 6000 Nano Kit (Agilent Technologies, USA). One of the TRIOL-treated monkeys was removed from RNA-seq analysis because the blood and brain RNA samples were poorly preserved due to depletion of dry ice during delivery to the sequencing company. One of the stage-A WBC samples from the HH+PROG group was also removed due to severe degradation of RNA.

For library construction, oligo-dT-coupled beads were first used to enrich poly-A+RNA molecules. First-strand cDNA synthesis was performed using random hexamers and Superscript II reverse transcriptase (Invitrogen, USA). Second-strand cDNA synthesis was performed using *E. coli* DNA PolI (Invitrogen, USA). A Qiaquick PCR purification kit (Qiagen, Germantown, MD, Germany) was used to purify the double-strand cDNA. cDNA was sheared with a nebulizer (Invitrogen, USA) into 100−500 bp fragments. The fragments were ligated to a Lumina PE adapter oligo mix after end repair and the addition of a 3' dA overhang. The 310−350 bp fragments were then collected by gel purification. After 15 cycles of PCR amplification, the libraries were subjected to paired-end sequencing using an HiSeq2500 (Illumina, USA)

platform. Finally, 357.58 Gb of transcriptome sequences were generated from the 91 samples for transcriptome analysis.

Measurement of intraneuronal calcium concentration ([Ca^{2+}]i)

The intracellular [Ca^{2+}]i of cultured primary rat cortical neurons was measured as per Minta et al. (1989). Briefly, after 8 d of *in vitro* culture on glass cover slips, primary cortical neurons were washed once with Hank's balanced saline solution (HBSS) loading buffer. We added 100 μL of HBSS loading buffer containing 5 μmol/L Fluo-3/AM (Molecular Probes, F1241, USA) dropwise onto the cover slips, which were then incubated with neurons for 45 min at room temperature before being washed twice with HBSS testing buffer to remove residual fluorescent dye. Neurons were then transferred to the perfusion chamber of an OctaFlow Perfusion System (ALA Scientific Instruments, Farmingdale, NY, USA), which was placed on the stage of a TCS SP2 laser scanning confocal microscope (Leica Microsystems, Mannheim, Germany). Using an excitation wavelength of 488 nm and emission wavelength of 526 nm, at least 10 randomly selected neurons on every slide were scanned continuously at 6 s intervals to determine real-time fluorescence intensity. Dynamic changes in the fluorescence intensity of every neuron were recorded automatically by the Leica confocal system. Changes in [Ca^{2+}]i (Δ[Ca^{2+}]i) were represented by changes in fluorescence intensity before and after drug administration and were expressed as: Δ[Ca^{2+}]i=(F (peak fluorescence intensity after administration) − F0 (baseline fluorescence intensity before administration))/F0×100%.

Statistical analysis of behavioral and neurological experiments

Statistical analyses were performed using SPSS 19.0 and graphs were drawn using GraphPad Prism 6.0. Most data, including brain water content and number of injured neurons, were analyzed using Student's *t*- test or one-way analysis of variance (ANOVA) with Dunnett's *post-hoc* test for multiple comparisons. Behavioral impairment scores were analyzed using the Kruskal-Wallis test. All data are represented as means±*SD* unless otherwise noted. *P*-values of less than 0.05 were considered statistically significant.

Source of genome reference and gene annotation

The reference genome and gene annotation of the cynomolgus monkey used in this study were originally generated by Yan et al. (2011) and downloaded from GigaDB (http://gigadb.org/dataset/100003). Specifically, the following files were downloaded:

Genome assembly: ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100003/CE.cns.all.fa.gz

Gene annotation (GFF): ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100003/CE.gff.gz

Gene annotation (CDS): ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100003/CE.cds.fa.gz

Gene annotation (PEP): ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100003/CE.pep.fa.gz

Transcriptome alignment and expression level measurement

The Hisat2 (v2.0.4) (Kim et al., 2015) package was employed to align transcriptome reads to the reference genome of cynomolgus monkey with default parameters. Gene expression was measured as RPKM (reads per kilobase of transcript per million mapped reads) using an in-house script (supplementary file "*count_reads_num. pl*"), and library sizes were normalized with the geometric method in DESeq2 (v1.10.1) (Love et al., 2014) (Supplementary Data). Genes expressed robustly (i.e., mean RPKM>1) in at least one sample group were used for principal component analysis (PCA).

Identification of differentially expressed genes (DEGs)

We applied a strict protocol to identify DEGs between groups of samples. Candidate DEGs were first identified using DESeq2 (v1.10.1) (Love et al., 2014), edgeR (v3.12.1) (Robinson et al., 2010), and Cuffdiff (v2.2.1) (Trapnell et al., 2013) separately. For DESeq2, the default method was used to normalize library sizes, the fitting type of dispersions to mean intensity was set as "*parametric*", and the method to test DEG significance was set as "*nbinomWaldTest*". For edgeR, the trimmed mean of M-values (TMM) normalization method was used to normalize library sizes, the quantile-adjusted conditional maximum likelihood (qCML) method was used to estimate dispersions, and the genewise exact test ("*exactTest*") was used to calculate the significance of DEGs. For Cuffdiff, the method for library size normalization was set as "*--geometric-norm*" and other parameters were set as default. P-values were adjusted to allow for multiple testing by the Benjamini-Hochberg False Discovery Rate (FDR) (Benjamini & Hochberg, 1995). A DEG was required to meet the following criteria: (1) at least two of the three DEG-detection software reported a FDR<0.05; (2) the lower quartile of expression level in the up-regulated group was greater than the upper quartile of expression level in the down-regulated group; and, (3) the mean RPKM was not less than 5 in the up-regulated group. Of note, for convenience, the FDR values shown for the genes mentioned in the Results section were all from edgeR.

To verify the robustness of our DEG-detection strategy and to calculate the false-positive rate, samples were randomly divided into pseudo groups. For WBC samples, three pseudo groups were generated, each consisting of two samples randomly chosen from stage A, stage D, and stage F, respectively. For brain samples, two pseudo groups were generated, each consisting of three samples randomly chosen from the HH and NN groups, respectively. The procedures for DEG detection described above were performed, and the numbers of DEGs between pseudo groups of the same tissue were considered as false positives. In contrast to the thousands of significant DEGs identified between real groups, no DEGs were identified between pseudo groups for WBCs and brain tissue, confirming the robustness of our strategy and reliability of our results.

Gene ontology annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation, and enrichment analysis

Gene ontology (GO) terms (Ashburner et al., 2000) for the cynomolgus monkeys were annotated according to their orthologous relationships with rhesus macaques and humans (Supplementary Table S5). KEGG pathways (Kanehisa & Goto, 2000) were annotated employing online KAAS (KEGG Automatic Annotation Server) tools (www.genome.jp/tools/kaas/) (Moriya et al., 2007) using all protein sequences of the cynomolgus monkeys as inputs with a BBH (bi-directional best hit) method (Supplementary Table S5).

Fisher's exact tests were employed to identify whether a list of genes (foreground genes) was enriched in specific GO terms or KEGG pathways, with comparisons of the number of foreground genes annotated to the specific GO/KEGG, number of foreground genes not annotated to the specific GO/KEGG, number of background genes (excluding foreground genes) annotated to the specific GO/KEGG, and number of background genes (excluding foreground genes) not annotated to the specific GO/KEGG. P- values were adjusted to allow for multiple testing by means of the Benjamini-Hochberg False Discovery Rate (Benjamini & Hochberg, 1995).

Co-expression module and hub gene identification

The WGCNA (v1.51) (Langfelder & Horvath, 2008) package in R (v3.2.3) was employed for co-expression network analyses using the percentage of transformed RPKM of all robustly expressed genes. Co-expression modules were identified using standard methods, with correlation between genes calculated by the Pearson method, and the distance for hierarchical clustering calculated by the Ward method (Ward Jr, 1963). Similar modules were merged by WGCNA with the parameter MEDissThres=0.25. The eigengene (vector presenting expression dynamics of a module between samples) and kME (correlation of the gene with corresponding module eigengene) in every module were also calculated by WGCNA. For every module, genes with kME>0.9 were considered hub genes.

Analysis of impact of drugs on transcriptome dynamics

For WBCs, the sequencing of the HH and drug-treated groups was not performed in the same batch. Hence, laboratory conditions, personnel differences, and reagent lots could have caused variations between the groups, i. e., batch effects (Leek et al., 2010). It should be noted that stage A samples from the PROG and TRIOL groups were collected prior to drug injection; therefore, the stage A DEGs between the HH and drug-treated groups were likely a reflection of batch effects rather than drug-specific effects, and were thus excluded from subsequent analyses on DEGs between HH and drug-treated groups in stages D and F.

For brains, we identified the DEGs between the NN, HH, and drug-treated groups in pairwise comparisons. Firstly, we focused on the impact of drug treatment on HH-responding

genes, which were DEGs between the NN and HH groups. If an HH-responding gene was a DEG between the HH and drug-treated groups but not a DEG between the NN and drug-treated groups (i. e., the expression level of the gene in the drug-treated group was not significantly different from that in the NN group but was significantly different from that in the HH group), the gene was considered to be a drug strongly responding gene (SRG). If an HH-responding gene was a DEG between the NN and drug-treated groups but not a DEG between the HH and drug-treated groups (i.e., the expression level of the gene in the drug-treated group was not significantly different from that in the HH group but was significantly different from that in the NN group), the gene was considered to be a drug non-responding gene (NRG). The other HH-responding genes with expression levels that fell between the drug-treated, NN, and HH groups were considered drug weakly responding genes (WRGs). If a gene was not significantly differentially expressed by HH in contrast to NN (i.e., not a DEG between the HH and NN groups) but was significantly differentially regulated by a drug in contrast to both NN and HH conditions, it was considered to be a drug-induced DEG. Even if a gene was significantly up- or down-regulated by HH, if that gene was further significantly up- or down-regulated but not restored by the drug, it was also considered to be a drug-induced DEG, not an NRG.

RESULTS

Monkeys develop severe brain injuries under acute HH conditions that can be significantly alleviated by PROG and TRIOL

Twenty-four healthy, adult male cynomolgus monkeys were classified into four groups (n=6): i. e., normobaric normoxia (NN), HH, HH+PROG, and HH+TRIOL. To characterize the behavioral and pathological changes caused by acute HH and to assess the drug effects, animals from the HH, HH+PROG, and HH+TRIOL groups were placed in a hypobaric chamber, which mimicked ascending altitudes from 320 to 7 500 m at a velocity of 2–3 m/s (HH group). Altitude simulation for each monkey started at 320 m (stage A) and was then sequentially suspended at 3 000 m (stage B), 4 500 m (stage C), and 6 000 m (stage D) for 50 min each. Altitude was then maintained at 7 500 m for 24 h (stage E) and then a further 24 h (stage F) for a total of 48 h (Figure 1A). During HH treatment, monkeys in the HH+PROG and HH+TRIOL groups were given either PROG or TRIOL, respectively (Figure 1B, C ; see Methods). In parallel, monkeys in the NN group were placed in a normobaric normoxia environment at an altitude of 320 m to serve as controls.

The impact of acute HH exposure on behavior was assessed 2 h after ascending to 7 500 m. In sharp contrast to the NN monkeys, which exhibited high activity levels with frequent walking, climbing, and eating, the HH monkeys lost their balance and lay prostrate with limited body movements, or even remained in lateral or dorsal recumbency during the entire behavioral assessment. All six monkeys in the HH

group displayed anorexia, vomiting, motor deficits, and ataxia. In contrast, monkeys in the HH+PROG and HH+TRIOL groups were able to maintain their balance and remain seated, with frequent head and forelimb movements. Several PROG or TRIOL-treated animals exhibited normal feeding behavior or short-distance movements (a few steps), although vomiting still occurred in some monkeys (two PROG-treated monkeys and four TRIOL-treated monkeys). We then employed behavioral deficit scoring to obtain a quantitative measure of skeletal muscle coordination by recording the movements of each monkey over 30 min (Supplementary Table S1) (Kito et al., 2001). All NN monkeys scored 0 (i. e., walked normally), whereas the HH monkeys scored between 14 (i. e., could not stand or sit for long) and 18 (i. e., no movement). PROG and TRIOL-treated HH monkeys scored between 10 (i. e., able to stand spontaneously) and 14; this was significantly lower than the scores of untreated HH monkeys (ANOVA P<0.05, Figure 1D), indicating that PROG and TRIOL were able to partially restore motor coordination in HH monkeys.

Cerebral edema is normally manifested as the increase of the water content in the brain tissue. We found that brain water content was significantly higher in the HH group (76.71%±0.30%) than in the NN group (76.16%±0.26%) (ANOVA P<0.01, Figure 1E). The proportion of brain water content was significantly reduced in both the HH+PROG and HH+TRIOL groups to 76.12%±0.46% and 76.28%±0.23%, respectively, similar to levels in the NN group (Figure 1E). Additionally, the pericapillary space was markedly wider in untreated HH brains than in other groups (Figure 1F), suggesting that blood-brain barrier (BBB) disruption and vasogenic edema, which are considered to cause cerebral edema in patients with HACE (Bärtsch & Swenson, 2013), had been induced by HH and reversed by drug administration. Nissl staining of the frontal cortex revealed many injured neurons in the HH group only. These cells were characterized by dark staining, condensed nuclei, shrunken cell bodies, or weak staining with irregular shapes (Figure 1G). The proportion of injured neurons in the HH group was 21.1%, significantly higher than that in the NN group (1.4%) (ANOVA P<10^{-3}). This ratio was restored to 7.5% in the HH+PROG group (ANOVA P <0.01) and 6.7% in the HH+TRIOL group (ANOVA P<0.01, Figure 1H; see Methods).

Taken together, these findings suggest that cynomolgus monkeys develop severe motor impairments and brain damage after acute exposure to HH environments, and this can be significantly mitigated by either PROG or TRIOL.

WBC dynamic transcriptome analysis reveals three major regulatory modules responding to acute HH

To systematically investigate the genetic regulatory networks that participate in the HH response, we sequenced and analyzed the transcriptome of WBCs obtained at each of the six HH stages, A–F (n=6, 36 WBC samples in total). Based on PCA of 12 198 robustly expressed genes, the degree of HH exposure (i. e., a function of altitude plus exposure time), rather than individual differences, appeared to be the main

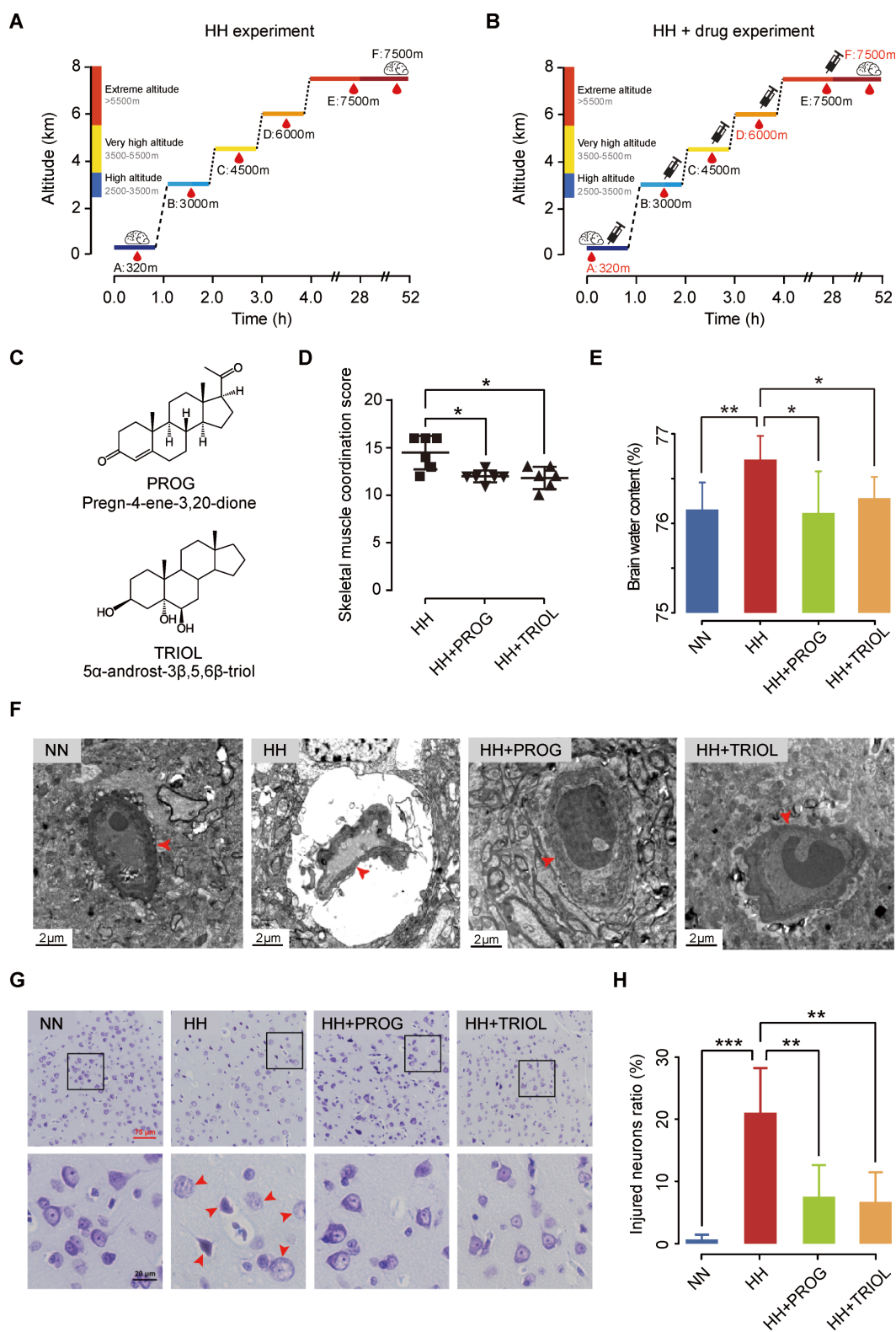**Figure 1 Acute HH-induced behavioral and cerebral impairments were attenuated by PROG and TRIOL treatments**

A: Experimental procedure for HH treatment of cynomolgus monkeys (HH group, *n*=6). Vertical bars along y-axis indicate medical definitions of high, very high, and extreme altitudes. Horizontal bars represent duration spent at each altitude. Blood drop and brain icons indicate altitudes and

time points at which blood and brain samples, respectively, were collected for RNA-seq analyses. B: Experimental procedure for drug treatment (HH+ PROG and HH+TRIOL groups, each *n*=6). Needle tubes indicate altitudes and time points at which drugs were injected; blood drop and brain icons indicate altitudes and time points at which blood and brain samples were collected, respectively. Drugs were injected after blood collection at each stage. Samples from stages highlighted in red were selected for RNA-seq analyses. C: Chemical structures of two neuroprotectants (PROG and TRIOL) used in this study. D: Mean skeletal muscle coordination scores for each experimental group, with higher scores representing a lower degree of coordination, assessed after 2 h at 7 500 m. Each group was compared with HH group (*: Kruskal-Wallis test *P*<0.05). Data are presented with error bars indicating means±*SD*, *n*=6 per group. E: Mean brain water content for each experimental group, measured as 1−dry/wet weight of left hemisphere. Each group was compared with HH group (*: one-way ANOVA *P*<0.05; **: *P*<0.01, HH *n*=5, each other group *n*=6). Error bars represent *SD*. F: Representative images showing ultrastructure of capillaries in frontal cortex of each experimental group, taken after 48 h at 7 500 m. Blood-brain barrier (BBB) disruption and vasogenic edema evident in HH group but not in other groups, as characterized by severely shrunken capillaries with fluid penetrating into pericapillary space. Red arrowheads indicate capillary boundary. Scale bar: 2 µm. G: Representative images showing Nissl staining of frontal cortex tissues of each experimental group. Neuronal injuries evident in HH group and markedly attenuated in drug-treated groups. Black squares in first row denote areas magnified in second row. Red arrowheads indicate injured cells with dark staining, condensed nuclei, shrunken cell bodies, or weak staining with irregular shapes. Red scale bar: 75 µm, black scale bar: 20 µm. H: Percentage of injured neurons in frontal cortex of each experimental group based on Nissl staining shown in panel G. Each group was compared with HH group (**: one-way ANOVA *P*≤0.01; ***: *P*≤0.001, NN *n*=5, all other groups *n*=4). Error bars represent *SD*.

factor affecting the genome-wide WBC gene expression patterns (Figure 2A). We then identified DEGs based on comparisons between any two of the six HH stages (see Methods) and obtained a total of 5 174 DEGs across all comparisons (ranging from 0–3 864 for different comparisons). As stage A (simulated altitude of 320 m) represents the NN condition, the number of DEGs between this and other stages reflects the degree of gene expression change in response to varying levels of HH exposure. The number of DEGs increased incrementally from stages B to D, reaching its highest at stage D. During stages E and F, the number decreased (Figure 2B), implying that the increased gene expression changes associated with HH were reversed in later stages, probably because of organismal acclimatization to hypoxia after 24 h of exposure to HH (Berglund, 1992).

Co-expression analysis was next used to identify co-expression modules associated with HH changes (see Methods). Eight modules were identified, ranging in size from 64 to 5 607 genes (Supplementary Figure S1A, B). Three modules were significantly enriched with DEGs (Fisher's exact test *P*<10^{-15}) and showed strong correlation between the expression pattern and progression of HH (ANOVA *P*<10^{-4}; Supplementary Table S2). Hence, these were considered the primary HH-responding modules (HH modules). The first HH module contained 3 114 genes, 2 123 (68%) of which were DEGs (Supplementary Table S2). This module (termed the 'impulse module') showed an 'impulse pattern', whereby the expression levels were monotonically increased or decreased in the early stages of HH (stages A–D), but recovered to a certain extent during the later stages (stages E–F) (Figure 2C). The second HH-responding module (termed the 'sustained module') contained 571 genes, 479 (84%) of which were DEGs (Supplementary Table S2). In this module, changes in expression levels of genes that responded to early-stage changes in altitude were sustained in the later stages of HH (Figure 2C). The third module (termed the 'late-responding module') contained 552 genes, 398 (72%) of which were DEGs (Supplementary Table S2). Genes in this

module showed no differential expression until the later stages of HH (Figure 2C).

Impulse responses of hypoxia/HIF-1 mediated innate immunity and inflammatory processes to acute HH
Genes in the impulse module were significantly implicated in several classical pathways critical to hypoxia response and inflammation (Figure 2D–E, and Supplementary Figure S1C). As noted in previous small mammal and human cell line studies, we confirmed that the HIF-1 signaling pathway also played a regulatory role in adaptive responses promoting optimal oxygen utilization (Semenza, 2009) and in inflammatory responses (Palazon et al., 2014) to hypoxia in *in vivo* primate experiments. Other innate immune and inflammatory pathways such as the NF-κB, Toll-like receptor, and FoxO signaling pathways, were also activated, probably forming a positive feedback loop with HIF-1 (Han et al., 2016). The inflammatory factor *IL1B* and chemokine *CXCL1*, which constitute links with HIF-1 and NF-κB signaling (Hatfield et al., 2010; Jung et al., 2003), were up-regulated by HH in the impulse module and are known to contribute to vascular impairment and vasogenic edema (Stamatovic et al., 2006), as observed in the brains of the HH-treated group (Figure 1F). Collectively, HIF-1/hypoxia pathways mediating the inflammatory response may contribute to pathogenesis in monkeys and were expressed in the impulse module during HH progression (Figure 2E). In contrast, the sustained module included genes mostly involved in adaptive immune responses (Figure 2D and Supplementary Figure S1C). Specifically, most key T cell markers, T cell receptors, and their specific kinases were consistently down-regulated by HH, suggesting overall suppression of the adaptive immune system in the HH environment.

Emergence of RBC-associated inflammatory processes in late stages
The late-responding module mainly contained genes involved in erythropoiesis and platelet activation (Figure 2D and Supplementary Figure S1C). Several genes encoding key

Figure 2 Transcriptomic dynamics of WBCs in response to acute HH.

A: Principal component analysis (PCA) using genes robustly expressed in at least one of six HH stages (A–F). Colors denote stages of HH at which WBC samples were collected, as described in Figure 1A. B: Numbers of differentially expressed genes (DEGs) between stage A and each of the other five stages. C: Expression patterns of DEGs in three HH-responding modules, shown as log2-transformed reads per kilobase of transcript per million mapped reads (RPKM) fold changes of DEGs in each stage compared to stage A. Red and light blue lines indicate expression patterns of individual DEGs up- and down-regulated by HH, respectively. Bold lines show mean expression levels of all up- or down-regulated DEGs, with error bars indicating SD. UP and DOWN indicate exact number of up- or down-regulated DEGs by HH in each module. D: Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in three HH-responding modules. Only significantly enriched pathways with a false discovery rate (FDR)-adjusted *P*-value of <0.05 were plotted. Dot color denotes *P*-value after FDR correction, and dot size denotes ratio of DEGs versus all expressed genes in each pathway. Different background colors denote classification of pathways. E: Interaction of pathways and symptoms according to reported articles. Blue color denotes pathways of impulse module; red color denotes pathways of late-responding module. Solid arrow line denotes reported direct interaction; dashed arrow line denotes reported associated relationship. F: Local co-expression network revealing genes in vitamin D receptor (VDR) complex as hub genes. Dark blue circles denote DEGs in impulse module; red circles denote genes in VDR complex and HIF1A as hub genes.

factors in erythropoiesis, including erythroid transcription factors (e.g., *GATA1* and *KLF1*), erythropoietin receptor *EPOR*, alpha and beta-globin genes (e.g., *HBA1*, *HBA2*, and *HBB*), hemoglobin stabilizing protein *AHSP*, key synthases of heme (e.g., *HMBS*, *ALAD*, and *UROS*), tropomodulin-1 *TMOD1*, and E3 ubiquitin-protein ligase *TRIM58*, were significantly up-regulated during stages E and F (Supplementary Figure S1D). This suggests that the enhanced expression of erythropoiesis genes was induced as an adaptive response after 28–52 h (stages E–F, Figure 1A) of HH exposure, which can regulate erythropoiesis and improve oxygen supply, thereby attenuating the impact of hypoxia in later stages. Key genes of blood coagulation and inflammation were also significantly up-regulated in late stages, including *GP1BA*, *F2R*, and other important procoagulant/antifibrinolytic factors (e.g., *VWF*, *SERPINE1*, and *F13A1*) that are associated with impairments of RBCs and vascular inflammation (Bergmeier et al., 2006; Gemmati et al., 2016; Sparkenbaugh & Pawlinski, 2013). In addition, the expression of the haptoglobin gene *HP*, a major scavenger of free hemoglobin, was also significantly up-regulated by HH (Supplementary Figure S1D), reflecting a homeostatic response to the increased release of free hemoglobin from impaired RBCs.

Intriguingly, in the late-responding module, we also found genes encoding critical markers of neutrophil degranulation, including neutrophil elastase (*ELANE*), myeloperoxidase (*MPO*), and cathepsin G (*CTSG*), the expression levels of which were up-regulated by HH (Supplementary Figure S1D). These activated neutrophils are tightly connected with BBB disruption, vascular inflammation, and vasogenic edema in injured brains (Ikegame et al., 2010; Liu et al., 2018), which could contribute to the development of the HACE-like symptoms observed in monkeys.

Vitamin D receptor (VDR) signaling is a novel key regulator involved in HH response

We next searched for intra-modular hub genes (i. e., genes with the highest degree of connectivity with other genes in the same module) in the three HH-responding modules to identify core regulators of the regulatory networks (see Methods). A total of 326 hub genes were identified (Supplementary Table S3), including those encoding key transcription factors, kinases, and receptors from the HIF-1 and FoxO signaling pathways (e.g., *STAT3*, *FOXO4*, and *PTEN* in the impulse module), T cell receptor signaling (e.g., *LCK* and *PTPN4* in the sustained module), and erythrocyte development and maintenance (e.g., *TMOD1*, *TRIM58*, and *KLF1* in the late-responding module).

Interestingly, our hub gene analysis also identified the vitamin D receptor (VDR) signaling pathway as an important but previously unreported pathway involved in the HH response. Specifically, genes encoding two key transcription factors, *VDR* and retinoid X receptor alpha (*RXRA*), and two co-activators, nuclear receptor coactivator 1 (*NCOA1*) and tripartite motif containing 24 (*TRIM24*), were identified as hub genes in the impulse module and were up-regulated at an early stage (Figure 2F). *VDR* and *RXRA* encode proteins that together form a heterodimeric complex that serves as a nuclear receptor for calcitriol, the hormonally active metabolite of vitamin D (Kato, 2000). Previous studies have revealed that VDR can inhibit the transcriptional activity of NF-κB (Chen et al., 2013b) and HIF-1 (Chung et al., 2009). VDR may also be up-regulated by Toll-like receptors and participate in the regulation of both the innate and adaptive immune systems (Adams & Hewison, 2008). These observations raise the possibility that the VDR pathway may function as a hub regulatory network that balances HIF-1 signaling and its inflammatory responses. Consistent with this, the vitamin D-binding protein in humans is associated with HH adaption in high-altitude dwelling native peoples (Ahmad et al., 2013), whereas VDR in mouse endothelial cells protects brains from hypoxia/reoxygenation-induced BBB disruption (Won et al., 2015).

Diverse effects of PROG and TRIOL on HH-induced WBC transcriptome dynamics

To evaluate the effects of two neuroprotective steroid agents on the blood cell transcriptome, we collected WBC samples for RNA-seq analysis from the HH+PROG and HH+TRIOL groups at stages D and F (Figure 1B), because stage D exhibited the most dynamic response to HH (Figure 2B) and stage F was the final stage of our analysis. WBC samples were also collected from the same animals at stage A, before drug injection, to serve as controls (Figure 1B). PCA revealed that neither PROG nor TRIOL treatment of HH animals had a significant impact on their WBC transcriptome dynamics, and the degree of HH exposure remained the primary differentiating factor between samples (Figure 3A).

To further quantify the potential effects of the two neuroprotective agents on specific genes and pathways, we identified agent-induced DEGs by comparing the gene expression between samples from agent-treated (HH+PROG or HH+TRIOL) and untreated HH groups at stages D and F. We identified 134 DEGs induced by PROG in stages D (70 DEGs) and F (72 DEGs) and found that these DEGs overlapped significantly with those from the late-responding module (Figure 3B and Supplementary Figure S2A, Fisher's exact test: $P<10^{-12}$). In particular, key erythrocyte-associated genes were among the DEGs with the largest expression changes (top DEGs) in stage F and significantly up-regulated by PROG, including the alpha and beta-globin genes (*HBA1*: FC=3.03, FDR=0.000 5; *HBA2*: FC=2.99, FDR=0.000 8; *HBB*: FC=3.40, FDR=0.000 1), heme synthases (*SLC25A39*: FC=2.10, FDR=0.008 3; *ALAS2*: FC=3.43, FDR=0.000 3), and bisphosphoglycerate mutase (*BPGM*: FC=2.33, FDR=0.026 6) (Figure 3C). Thus, although PROG appeared to have a relatively subtle effect on the WBC transcriptomes, we speculate that this steroid may function specifically to enhance erythropoiesis and elevate oxygen transport in the blood, thereby helping to relieve the symptoms caused by HH exposure.
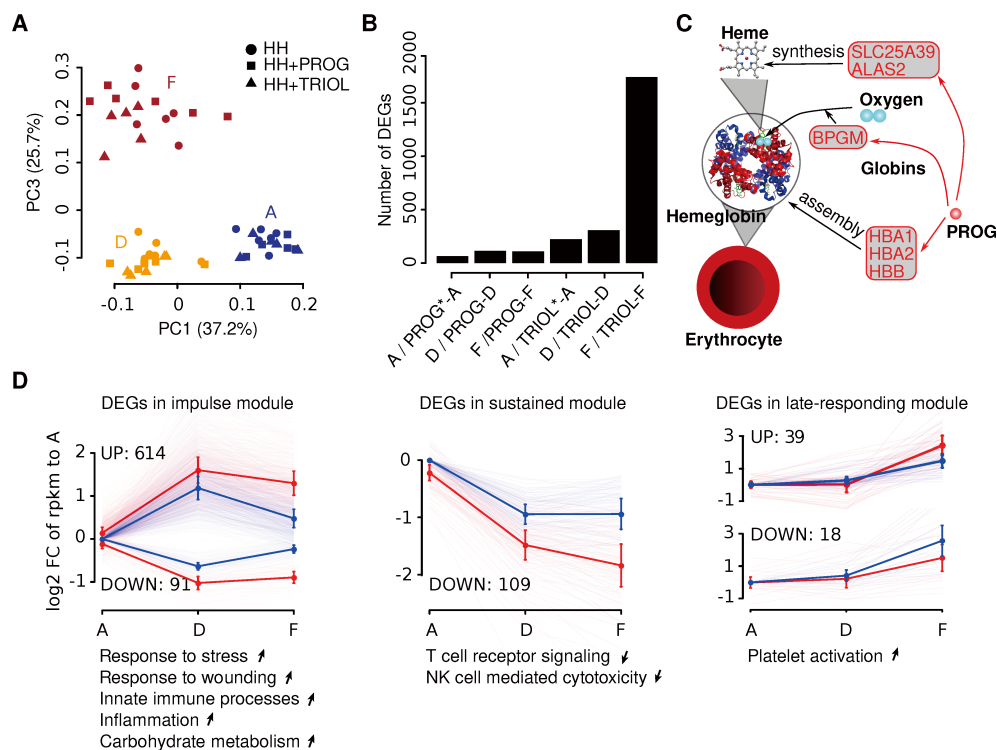
Figure 3 Regulation of WBC transcriptomic dynamics by PROG and TRIOL

A: Principal component analysis (PCA) using genes robustly expressed in at least one of nine sample groups (i.e., no-drug treated group, PROG-treated group, and TRIOL-treated group in stages A, D, and F, respectively). Colors denote HH stages at which WBC samples were collected during HH+drug experiments, as described in Figure 1B, whereas shapes denote treatments. B: Number of differentially expressed genes (DEGs) between drug-treated and untreated groups in each HH stage studied. Stage A samples from both PROG and TRIOL groups were collected prior to drug injection; thus, DEGs of A/PROG*-A and A/TRIOL*-A reflect random variations between two groups of experimental monkeys rather than drug-specific effects. C: Erythrocyte-associated DEGs up-regulated by PROG. Red color and arrows denote genes up-regulated by PROG. D: Expression patterns and functions of TRIOL-induced DEGs in three HH-responding modules, shown as log2-transformed reads per kilobase of transcript per million mapped reads (RPKM) fold changes of DEGs in each sample group compared to stage A samples without TRIOL treatment. Light red and blue lines indicate expression patterns of individual DEGs in TRIOL-treated and untreated groups, respectively. Bold lines show mean expression levels of all TRIOL up- or down-regulated DEGs in each group, with error bars representing *SD*. UP and DOWN indicate exact number of up- or down-regulated DEGs by TRIOL in each module. Text below each module outlines functions associated with TRIOL-induced DEGs, with up and down arrows denoting gene-associated functions up- or down-regulated by TRIOL, respectively.

Compared with PROG, TRIOL had significantly broader effects on the WBC transcriptome (Figure 3B), inducing 1 676 DEGs during stages D (195 DEGs) and F (1 605 DEGs). Interestingly, these TRIOL-induced DEGs overlapped significantly with HH-induced DEGs in all three HH-responding modules (Supplementary Figure S2A). Functional enrichment analysis revealed that TRIOL significantly up-regulated HIF1/hypoxia and RBC-associated inflammation, and down-regulated T cell receptor signaling during stage F (Figure 3D and Supplementary Figure S2B). Thus, in contrast to PROG, which had a specific function in erythropoiesis, TRIOL treatment altered the expression of a broader variety of genes in WBCs in terms of HH-related functions.

Profound transcriptomic alterations underlie pathological effects of HH-induced brain injuries

To investigate the brain gene regulatory network underlying the acute HH response, and the effects of PROG and TRIOL, we performed RNA-seq on frontal cortex samples (brains) from all monkeys after HH treatment (Figure 1A, B). Based on PCA of 14 200 robustly expressed genes, the brain samples could be clearly separated into the three groups: NN, HH, and drug treatment. The HH and NN groups could be separated at both the PC2 and PC3 level, as was performed for the WBC samples, indicating that HH was the main factor affecting the global brain gene expression pattern for monkeys without drug treatment (Figure 4A). Moreover, both agent-treated groups clustered tightly together, but separately from the HH group at the PC2 level (34.8% of total variance) and from the NN group at the PC3 level (23.8% of total variance) (Figure 4A), suggesting marked effects of both neuroprotective steroid agents on the brain transcriptome.

We identified 2 992 DEGs between the NN and HH groups

Figure 4 Transcriptomic changes in frontal cortex in response to acute HH, PROG, and TRIOL

A: Principal component analysis (PCA) using genes robustly expressed in at least one of four sample groups (normobaric normoxia [NN], HH, HH+ PROG, and HH+TRIOL). Colors denote sample group. B: Expression changes of HH-induced differentially expressed genes (DEGs) after PROG or TRIOL treatment, shown as log2-transformed reads per kilobase of transcript per million mapped reads (RPKM) fold changes of DEGs in each sample group compared to NN group. HH-induced DEGs are categorized into strongly responsive genes (SRGs), weakly responsive genes (WRGs), and non-responsive genes (NRGs) according to their degree of expression level recovery after drug treatment. Red and light blue lines denote expression patterns of individual DEGs up- and down-regulated by HH, respectively. Bold lines show mean expression levels of all HH up- or down-regulated DEGs in each group, with error bars representing SD. Percentages of SRGs and WRGs versus all HH-induced DEGs are presented above each plot. C: Expression changes of drug DEGs after PROG or TRIOL treatment, shown as log2-transformed RPKM fold changes of DEGs in each sample group compared to NN group. Red and light blue lines denote expression patterns of individual DEGs up- or down-regulated by drugs, respectively. Bold lines show mean expression levels of all DEGs up- or down-regulated by drugs in each group, with error bars representing SD. D: Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in HH-induced or drug-recovered DEGs. Only significantly enriched pathways with a false discovery rate (FDR)-adjusted P-value of <0.05 are plotted. In HH DEGs, SRGs, WRGs, and NRGs, red and blue dots denote pathways enriched in DEGs up- and down-regulated by HH, respectively. In drug DEGs (PROG DEGs and TRIOL DEGs), red and blue dots denote pathways enriched in DEGs up- and down-regulated by drugs, respectively. Color intensity of dot indicates level of significance, and size of dot denotes ratio of DEGs versus all expressed genes in each pathway. E: Postulated regulation of excitatory glutamate signaling by TRIOL. Rectangles represent proteins of TRIOL DEGs involved in glutamate signaling pathway. Blue colors denote down-regulation by HH. Gene symbols of DEGs associated with each protein in the pathway are listed below. F: Time-course showing intraneuronal calcium $[Ca^{2+}]_i$ in primary cultured rat cortical neurons stimulated by glutamate with different doses of TRIOL. Glutamate stimulation was added at the fiftieth second in the experiment.

(see Methods), with 1 446 up- regulated and 1 546 down-regulated by HH (Supplementary Table S4), reflecting considerable regulatory and functional rewiring in brainexpressed genes in response to acute HH exposure. DEGs showing the largest expression fold changes (top 1% of all DEGs) comprised the hemoglobins (carrier of oxygen) HBA1 (FC=10.69, FDR=0.000 0) and HBA2 (FC=4.12, FDR=0.000 2), and several angiogenesis-related genes, including adrenomedullin ADM (FC=13.58, FDR=0.000 0), vascular endothelial growth factor VEGFA (FC=5.48, FDR=0.000 0), and apelin APLN (FC=3.29, FDR=0.000 0), which were all up-regulated in HH brains. It is worth noting that many tightly associated DEGs were also significantly up-regulated by HH, including VEGFB (F=1.58, FDR=0.000 0), receptor of vascular endothelial growth factor FLT1 (FC=2.62, FDR=0.000 0), and receptor of adrenomedullin RAMP2 (FC=1.45, FDR=0.006 2). This strongly suggests that an adaptive response to HH, by increasing local tissue oxygenation and increasing oxygen delivery, occurred in monkey brains after 52 h of HH exposure. However, VEGF expression can also promote BBB leakage (Schoch et al., 2002), as observed in all monkeys exposed to HH (Figure 1F). Of note, CARTPT , a gene that encodes neuropeptides controlling feeding (Lambert et al., 1998), locomotor activity (Kimmel et al., 2002), and stress response (Rogge et al., 2008), also emerged as a top DEG (FC=6.88, FDR=0.000 3) between NN and HH groups, and its up-regulation in HH brains likely played a role in inducing the symptoms of anorexia and motor deficits in HH monkeys.

As with WBC transcriptome analysis, functional enrichment analyses revealed that HIF-1 signaling was significantly enriched by DEGs up-regulated in HH brains (Figure 4D), including EPAS1 (FC=1.37, FDR=0.028 5) and EGLN1 (FC=1.48, FDR=0.004 7), two genes strongly associated with high-altitude adaptation in highland populations (Simonson et al., 2010; Yi et al., 2010), and HIF3A (FC=4.36, FDR=0.000 0), an important regulator of HIF1A and EPAS1 involved in adaptive responses to hypoxia (Ravenna et al., 2015). Moreover, given that the top DEGs (e.g., ADM, VEGF, APLN, HBA1, and HBA2) are regulated by HIF-1 signaling (Chen et al., 2012; Eyries et al., 2008; Ramakrishnan et al., 2014; Saha et al., 2014), this suggests that HIF-1 signaling also plays a main regulatory role in adaptive responses to acute HH in brains. Additionally, genes involved in the complement and coagulation cascades, and in leukocyte transendothelial migration, were up-regulated in HH-exposed brains (Figure 4D and Supplementary Figure S3A) and WBCs (Figure 2D), implying a link between peripheral inflammation and neuroinflammation in the CNS.

As a consequence of the decrease in oxygen supply under HH conditions, several pathways involved in aerobic respiration were found to be down-regulated in HH brains, including oxidative phosphorylation, respiratory electron transport chain, tricarboxylic acid (TCA) cycle, and pyruvate metabolism. On the other hand, pathways involved in anaerobic respiration were up-regulated, including glycolysis

and mTOR signaling (Figure 4D and Supplementary Figure S3A). These findings suggest that, regardless of the adaptive response by stimulating angiogenesis and increasing oxygen delivery, the monkey brains suffered an energy deficiency, which may have induced autophagy and neuronal injuries through mTOR signaling (Dunlop & Tee, 2014).

Protein misfolding and endoplasmic reticulum (ER) stress are involved in many types of brain injury and neurodegenerative diseases (Giffard et al., 2004; Soto & Estrada, 2008). Indeed, we found that many key genes involved in the protein folding process were down-regulated by HH (Supplementary Figure S3A), including those encoding members of the HSP40 family, prefoldins, and FK506 binding proteins (Supplementary Table S4). Members of the HSP40 family are crucial partners for HSP70 chaperones (Qiu et al., 2006), and are important for hypoxic tolerance (Jain et al., 2014) and injury resistance (Hasegawa et al., 2018) in the CNS. Thus, their down-regulation may exacerbate HH-induced brain injury.

Taken together, our results revealed profound neural gene expression alterations after exposure to acute HH, and rewiring of genetic pathways, which may contribute to an adaptive response to acute HH and underlie the pathological effects of the associated brain injuries.

Brain transcriptomic changes induced by HH are reversed by treatment with PROG and TRIOL
We then assessed how PROG and TRIOL affect the expression of HH-altered genes (see Methods). Of all HH-altered genes, the expression levels were restored to normal (defined by reference to the NN group) in 50.67% and 48.09% of cases by PROG and TRIOL treatment, respectively, i. e., 'strongly responsive genes' (SRGs) (Figure 4B). Furthermore, 31.52% and 29.45% of HH-altered genes were partially restored by PROG and TRIOL treatment, respectively, i. e., 'weakly responsive genes' (WRGs) (Figure 4B). In addition, the expression levels of 17.15% and 20.82% of HH-altered genes were unaffected by PROG and TRIOL treatment, respectively, i.e., 'non-responsive genes' (NRGs) (Figure 4B). These results indicate that the abnormal expression levels of most HH-induced DEGs were completely or partially restored by PROG (82.19%) or TRIOL (77.54%) treatment. It should be noted that the SRGs and WRGs associated with the two drug treatments largely overlapped with each other (Supplementary Figure S3B), suggesting that PROG and TRIOL have broadly similar effects on reversing HH-induced transcriptomic changes in the brain.

PROG and TRIOL SRGs were particularly enriched in components of the key pathways of aerobic respiration (Figure 4D and Supplementary Figure S3A), suggesting an improvement in aerobic energy supply in response to PROG and TRIOL treatment. Specifically, of the 44 genes involved in oxidative phosphorylation and down-regulated by HH, 25 (56.8%) were PROG SRGs and 39 (88.6%) were TRIOL SRGs (Fisher's exact test, P=0.002). This implies that TRIOL may be more effective at improving the aerobic energy supply

than PROG. The abnormal HH-induced expression of genes with roles in protein folding was also restored by both PROG and TRIOL treatment (Figure 4D and Supplementary Figure S3A), which may relieve neuronal injuries induced by protein folding and ER stress. In addition, the up-regulation of the behavior-related gene *CARTPT* by HH was also attenuated by both drugs to a certain degree (Supplementary Table S4), consistent with the alleviation of anorexia and motor deficits in drug-treated monkeys.

Of note, the up-regulations of several top DEGs involved in angiogenesis by HH were not recovered by PROG or TRIOL, including *ADM*, *VWF*, and *VEGFA*, suggesting that angiogenesis was also active in drug-treated groups (Supplementary Table S4). Intriguingly, PROG but not TRIOL further up-regulated several key genes involved in angiogenesis and oxygen delivery, including *VEGFA* (HH-vs-PROG: FC=1.23, FDR=0.032 3), its receptor *FLT1* (HH-vs-PROG: FC=1.23, FDR=0.032 3), and hemoglobins *HBB* (HH-vs-PROG: FC=2.33, FDR=0.002 4) and *HBA2* (HH-vs-PROG: FC=1.97, FDR=0.000 0) (Supplementary Table S4). These results further emphasize the potential role of PROG in relieving HH-induced symptoms by elevating oxygen transport in both WBCs and brains.

Gene regulatory mechanisms by which TRIOL suppresses glutamate-mediated excitotoxicity

In the brain transcriptome, we detected many genes that were specifically rewired by either PROG or TRIOL, even though their expression levels were not altered by HH treatment (see Methods, Figure 4C). These included 380 and 743 genes whose expression levels were induced by PROG and TRIOL, respectively. Intriguingly, among them, we found that TRIOL, but not PROG, significantly induced the enrichment of DEGs in many key excitatory neuronal signaling pathways (Figure 4D and Supplementary Figure S3A), including the glutamatergic synapse, L-glutamate import, glutamate receptor, and calcium signaling pathways. This suggests that TRIOL may have a specific effect on the regulation of glutamate-induced excitotoxicity. Glutamate induces excitotoxicity by promoting the excessive influx of intracellular calcium $[Ca^{2+}]$ through ionotropic glutamate receptors, particularly through the N-methyl-d-aspartate (NMDA) receptor. This is the main cause of hypoxia-induced neuronal death in many hypoxia-induced brain injuries (Wroge et al., 2012). Many key genes involved in this process were down-regulated by TRIOL, including NMDA receptors *GRIN2A* (FC=1.79, FDR=0.000 0)and *GRIN2B* (FC=2.23, FDR=0.000 3), vesicular glutamate transporter (VGLUT) *SLC17A7* (FC=1.63, FDR=0.000 0), excitatory amino acid transporter (EAAT) *SLC1A2* (FC=1.60, FDR=0.018 4), group I metabotropic glutamate receptor *GRM5* (FC=1.48, FDR=0.030 4), and its downstream adenylate cyclase family members *ADCY1* (FC=1.52, FDR=0.029 1), *ADCY3* (FC=1.39, FDR 0.009 8), *ADCY5* (FC=1.55, FDR 0.003 8), and *ADCY9* (FC=1.59, FDR=0.000 2) (Figure 4E and Supplementary Table S4).

Glutamate-induced intracellular calcium ($[Ca^{2+}]_i$) overload is part of the signature indicating excitotoxicity in neurons (Scannevin & Huganir, 2000). To further examine whether TRIOL can indeed reduce neuronal excitotoxicity by directly suppressing the intracellular accumulation of calcium, we measured glutamate-induced $[Ca^{2+}]i$ in primary rat cortical neurons with and without TRIOL treatment. We observed that stimulation with 100 μmol/L glutamate alone dramatically increased $[Ca^{2+}]_i$ in neurons within 50 s , whereas supplementation with TRIOL decreased these glutamate-induced effects in a dose-dependent fashion (Figure 4F). Taken together, our results suggest that TRIOL may suppress excitotoxicity via the indirect transcriptional regulation of key glutamate signaling-related genes through direct modulation of membrane permeability for calcium.

DISCUSSION

In this study, we demonstrated that cynomolgus monkeys exposed to acute HH exhibited various symptoms, including anorexia, vomiting, motor deficits, and ataxia, as observed in human patients suffering acute high-altitude diseases and which are mostly attributed to pathological changes in the brain (Wilson et al., 2009). The pathophysiological diagnoses in the brain suggest that cynomolgus monkeys developed severe cerebral edema, similar to that found in human HACE patients (Hackett & Roach, 2004). Hence, these data suggest cynomolgus monkeys can be used as non-human primate models for studying the disorders of oxygen homeostasis. Moreover, the genetic regulatory networks shown in this study to be involved in the development of hypoxia-induced brain malfunctions provide valuable insights into the relationship between the mechanism of oxygen homeostasis regulation and related diseases. Our study confirmed the fundamental role of several classical pathways, such as HIF-1 signaling, in the organism level response to hypoxia. Moreover, we also revealed that several molecules, such as VDR, serve as central connectors of different acute HH response pathways; these have been largely ignored in previous studies using human cell lines or mice as experimental models.

The dynamic transcriptome of WBCs from non-human primates may help to improve our understanding of the progression of HACE. First of all, most genes were categorized as belonging to the impulse module, including those involved in HIF-1 signaling; the expression of these genes was stimulated during early HH stages but recovered in late stages, suggesting that hypoxia is indeed the primary cause of injury under acute HH conditions. The hub genes in the impulse module may account for the subsequent expression changes in genes in the late stages of HH and may be responsible for the fundamental pathophysiological changes involved in acclimatization or injury. For example, RBC-associated inflammation is a novel pathway in late stages and may interact with the progress of HACE. Our transcriptomic analyses combining WBCs and brains also highlighted a complex organismal response to acute HH, with active crosstalk between the acclimatization that maintains

oxygen homeostasis and adverse pathological feedback caused by inflammation. Specifically, in addition to the activation of the HIF-1 signaling pathway, consistent with previous reports (Semenza, 2009; Wilson et al., 2009), we also found that HH may induce a hierarchy of transcriptional programs of both innate and adaptive immunity, while progressively promoting inflammation in both the blood and brain.

The development of treatments for brain injury associated with HH, such as HACE, remains a formidable challenge. Our pharmacogenomic studies reported here provide preliminary but nevertheless valuable insights into the feasibility of a generalized neuroprotection strategy. Further, they present possible drug candidates, PROG and TRIOL, which both alleviated acute HH-induced brain injuries in our primate model. PROG is reported to be neuroprotective by stimulating breathing, increasing oxygen-carrying capacity, and modulating the GABAa receptor to suppress excitotoxicity (Singh & Su, 2013), whereas our results suggested a role in stimulating erythropoiesis and increasing oxygen delivery. Medroxyprogesterone (MPA), a synthetic analogue of PROG without neuroprotectant properties (Singh & Su, 2013), is used to treat chronic mountain sickness by stimulating breathing and increasing blood oxygen levels (Wright et al., 2004). However, MPA has no significant prophylactic effect for acute mountain sickness (AMS) (Wright et al., 2004), which further emphasizes the importance of neuroprotection in treating HACE. It is intriguing that the two neuroprotectants, PROG and TRIOL, appear to have different modes of action and act through distinct functional mechanisms to rewire the expression of genes in the blood and brain. Our previous study reported that TRIOL can protect primary cortical neurons from hypoxia/reoxygenation injury by maintaining mitochondrial functions (Chen et al., 2013a). Here, we further revealed that the neuroprotective effect of TRIOL may be mediated by attenuation of glutamate-induced excitotoxicity.

In conclusion, our research extends the transcriptomic analysis of acute HH responses at an integrated spatial and temporal level in non-human primates. We presented a kinetic transcriptomic analysis of WBCs and the transcriptomic profile of brains after exposure to acute HH. Our data revealed thousands of genes regulated by acute HH and categorized to different co-expression modules depending on their spatial and temporal dimensions. Our study also revealed that neuroprotective steroids such as PROG and TRIOL have significant recovery effects on acute HH-induced behavioral deficits and brain damage, suggesting the potential of using novel neuroprotectants as therapeutic drugs for acute high-altitude diseases. However, we also acknowledge that there are ~25 million years of evolutionary divergence between cynomolgus monkeys and humans (Yan et al., 2011), and some routine physiological data such as arterial blood pressure, arterial oxygenation, and ventilation were not collected in this study due to the limitation of experimental conditions. Thus, we cannot rule out that the pathological and gene regulatory changes observed in the cynomolgus

monkeys may differ from that in humans suffering HH-induced disorders. Future studies with relevant physiological measurements could help better understand how animals respond to severe hypoxia, and how neuroprotective steroids function to protect HH-induced brain damage.

DATA AVAILABILITY

The underlying sequence data used in this study were deposited with the National Center for Biotechnology Information (NCBI) under accession No. PRJNA431946 and in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession code CNP0000293. The DEGs are shown in Supplementary Tables S3 and S4.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

G.M.Y., G.J.Z., and Y.Q.G. designed the overall project; P.Z., Q.Y.L., W.Y., X. Y.Z., B.Z.L., L.X.S., and B.T.Q. carried out the transcriptome analyses; W.Y., Y.X.G., J.S.C., B.Z.L., W.B.Z., Z.B.Y., G.X., F.Y.L., and Z.W.T replicated animal models using the hypobaric chamber; B.Z.L., J.S.C., W.B.Z., M.Y., and Y.Q.W. prepared the samples for transcriptomic and pharmacodynamic testing; Q. Y. L., P. Z., and C. X. G. managed library construction and sequencing; B.Z.L., W.Y., J.S.C., and Y.L. performed the ultrastructural and histological assays; S. Z. L., J. X. Z., and Y. J. H. provided the TRIOL and pharmaceutical preparations; W.Y., J.S.C., and F.Y.L. handled the shipment and management of animals; J. W. and H. M. Y. contributed reagents/materials/analysis tools; W.Y., P.Z., Q.Y.L., B.Z.L., and L.X.S. drafted the manuscript; G.J.Z. and D.N.C. revised the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Adams JS, Hewison M. 2008. Unexpected actions of vitamin D: new perspectives on the regulation of innate and adaptive immunity. *Nature Clinical Practice Endocrinology and Metabolism*, **4**(2): 80−90.

Ahmad Y, Sharma NK, Garg I, Ahmad MF, Sharma M, Bhargava K. 2013. An insight into the changes in human plasma proteome on adaptation to hypobaric hypoxia. *PLoS One*, **8**(7): e67548.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics*, **25**(1): 25−29.

Bärtsch P, Swenson ER. 2013. Acute high-altitude illnesses. *The New England Journal of Medicine*, **368**(24): 2294−2302.

Basnyat B, Murdoch DR. 2003. High-altitude illness. *The Lancet*, **361**(9373): 1967−1974.

Basu M, Pal K, Prasad R, Malhotra AS, Rao KS, Sawhney RC. 1997. Pituitary, gonadal and adrenal hormones after prolonged residence at extreme altitude in man. *International Journal of Andrology*, **20**(3): 153−158.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**(1): 289−300.

Berglund B. 1992. High-altitude training. *Sports Medicine*, **14**(5): 289−303.

Bergmeier W, Piffath CL, Goerge T, Cifuni SM, Ruggeri ZM, Ware J, Wagner DD. 2006. The role of platelet adhesion receptor GPIbα far exceeds that of its main ligand, von Willebrand factor, in arterial thrombosis. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(45): 16900−16905.

Chen J, Leng T, Chen W, Yan M, Yin W, Huang Y, Lin S, Duan D, Lin J, Wu G, Zhang J, Yan G. 2013a. A Synthetic steroid 5α -androst-3β, 5, 6β -triol blocks hypoxia/reoxygenation-induced neuronal injuries via protection of mitochondrial function. *Steroids*, **78**(10): 996−1002.

Chen L, Qiu J-H, Zhang L-L, Luo XD. 2012. Adrenomedullin promotes human endothelial cell proliferation via HIF-1α. *Molecular and Cellular Biochemistry*, **365**(1–2): 263−273.

Chen Y, Zhang J, Ge X, Du J, Deb DK, Li YC. 2013b. Vitamin D receptor inhibits nuclear factor κB activation by interacting with IκB kinase β protein. *The Journal of Biological Chemistry*, **288**(27): 19450−19458.

Chung I, Han G, Seshadri M, Gillard BM, Yu WD, Foster BA, Trump DL, Johnson CS. 2009. Role of vitamin D receptor in the antiproliferative effects of calcitriol in tumor-derived endothelial cells and tumor angiogenesis in vivo. *Cancer Research*, **69**(3): 967−975.

Dunlop EA, Tee AR. 2014. mTOR and autophagy: a dynamic relationship governed by nutrients and energy. *Seminars in Cell & Developmental Biology*, **36**(5): 121−129.

Eyries M, Siegfried G, Ciumas M, Montagne K, Agrapart M, Lebrin F, Soubrier F. 2008. Hypoxia-induced apelin expression regulates endothelial cell proliferation and regenerative angiogenesis. *Circulation Research*, **103**(4): 432−440.

Gemmati D, Vigliano M, Burini F, Mari R, El Mohsein HH, Parmeggiani F, Serino ML. 2016. Coagulation factor XIIIA (F13A1): novel perspectives in treatment and pharmacogenetics. *Current Pharmaceutical Design*, **22**(11): 1449−1459.

Giffard RG, Xu L, Zhao H, Carrico W, Ouyang Y, Qiao Y, Sapolsky R, Steinberg G, Hu B, Yenari MA. 2004. Chaperones, protein aggregation, and brain protection from hypoxic/ischemic injury. *Journal of Experimental Biology*, **207**(18): 3213−3220.

Guo P, Luo H, Fan Y, Luo Y, Zhou Q. 2013. Establishment and evaluation of an experimental animal model of high altitude cerebral edema. *Neuroscience Letters*, **547**: 82−86.

Hackett PH, Roach RC. 2004. High altitude cerebral edema. *High Altitude Medicine & Biology*, **5**(2): 136−146.

Han S, Xu W, Wang Z, Qi X, Wang Y, Ni Y, Shen H, Hu Q, Han W. 2016. Crosstalk between the HIF-1 and Toll-like receptor/nuclear factor- κB pathways in the oral squamous cell carcinoma microenvironment. *Oncotarget*, **7**(25): 37773−37789.

Hasegawa T, Yoshida S, Sugeno N, Kobayashi J, Aoki M. 2018. DnaJ/Hsp40 family and Parkinson's disease. *Frontiers in Neuroscience*, **11**: 743.

Hatfield KJ, Bedringsaas SL, Ryningen A, Gjertsen BT, Bruserud O. 2010. Hypoxia increases HIF-1α expression and constitutive cytokine release by primary human acute myeloid leukaemia cells. *European Cytokine Network*, **21**(3): 154−164.

Hu H, Zhou Y, Leng T, Liu A, Wang Y, You X, Chen J, Tang L, Chen W, Qiu P, Yin W, Huang Y, Zhang J, Wang L, Sang H, Yan G. 2014. The major cholesterol metabolite cholestane-3β, 5α, 6β - triol functions as an endogenous neuroprotectant. *Journal of Neuroscience*, **34**(34): 11426−11438.

Huang X, Zhou Y, Zhao T, Han X, Qiao M, Ding X, Li D, Wu L, Wu K, Zhu LL, Fan M. 2015. A method for establishing the high-altitude cerebral edema (HACE) model by acute hypobaric hypoxia in adult mice. *Journal of Neuroscience Methods*, **245**: 178−181.

Ikegame Y, Yamashita K, Hayashi S, Yoshimura S, Nakashima S, Iwama T. 2010. Neutrophil elastase inhibitor prevents ischemic brain damage via reduction of vasogenic edema. *Hypertension Research*, **33**(7): 703−707.

Imray C, Wright A, Subudhi A, Roach R. 2010. Acute mountain sickness: pathophysiology, prevention, and treatment. *Progress in Cardiovascular Diseases*, **52**(6): 467−484.

Jain K, Suryakumar G, Ganju L, Singh SB. 2014. Differential hypoxic tolerance is mediated by activation of heat shock response and nitric oxide pathway. *Cell Stress Chaperones*, **19**(6): 801−812.

Jung YJ, Isaacs JS, Lee S, Trepel J, Neckers L. 2003. IL-1β mediated up-regulation of HIF-1α via an NFkB/COX-2 pathway identifies HIF-1 as a critical link between inflammation and oncogenesis. *The FASEB Journal*, **17**(14): 2115−2117.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **28**(1): 27−30.

Kato S. 2000. The function of vitamin D receptor in vitamin D action. *The Journal Biochemistry*, **127**(5): 717−722.

Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biology*, **8**(6): e1000412.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**(4): 357−360.

Kimmel HL, Thim L, Kuhar MJ. 2002. Activity of various CART peptides in changing locomotor activity in the rat. *Neuropeptides*, **36**(1): 9−12.

Kito G, Nishimura A, Susumu T, Nagata R, Kuge Y, Yokota C, Minematsu K. 2001. Experimental thromboembolic stroke in cynomolgus monkey. *Journal Neuroscience Methods*, **105**(1): 45−53.

Lambert PD, Couceyro PR, Mcgirr KM, Dall Vechia SE, Smith Y, Kuhar MJ. 1998. CART peptides in the central control of feeding and interactions with neuropeptide Y. *Synapse*, **29**(4): 293−298.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**(1): 559.

Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. 2010. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*, **11**(10): 733−739.

Lissa V-A, Bruno M, Suzana H-H. 2013. Different scaling of white matter volume, cortical connectivity, and gyrification across rodent and primate brains. *Frontiers in Neuroanatomy*, **7**(3).

Liu YW, Li S, Dai SS. 2018. Neutrophils in traumatic brain injury (TBI):

friend or foe?. *Journal of Neuroinflammation*, **15**(1): 146.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12): 550.

Martin D, Windsor J. 2008. From mountain to bedside: understanding the clinical relevance of human acclimatisation to high-altitude hypoxia. *Postgraduate Medical Journal*, **84**(998): 622−627.

Minta A, Kao JP, Tsien RY. 1989. Fluorescent indicators for cytosolic calcium based on rhodamine and fluorescein chromophores. *The Journal of Biological Chemistry*, **264**(14): 8171−8178.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**(suppl_2): W182−W185.

Nei M, Xu P, Glazko G. 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(5): 2497−2502.

Palazon A, Goldrath AW, Nizet V, Johnson RS. 2014. HIF transcription factors, inflammation, and immunity. *Immunity*, **41**(4): 518−528.

Qiu XB, Shao YM, Miao S, Wang L. 2006. The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. *Cellular and Molecular Life Sciences*, **63**(22): 2560−2570.

Rabinstein AA. 2006. Treatment of cerebral edema. *The Neurologist*, **12**(2): 59−73.

Ramakrishnan S, Anand V, Roy S. 2014. Vascular endothelial growth factor signaling in hypoxia and inflammation. *Journal of Neuroimmune Pharmacology*, **9**(2): 142−160.

Ravenna L, Salvatori L, Russo MA. 2015. HIF3a: the little we know. *The FEBS Journal*, **283**(6): 993−1003.

Robinson MD, Mccarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**(1): 139−140.

Rogge G, Jones D, Hubert GW, Lin Y, Kuhar MJ. 2008. CART peptides: regulators of body weight, reward and other functions. *Nature Reviews Neuroscience*, **9**(10): 747−758.

Saha D, Patgaonkar M, Shroff A, Ayyar K, Bashir T, Reddy KVR. 2014. Hemoglobin expression in nonerythroid cells: novel or ubiquitous?. *International Journal of Inflammation*, **2014**.

Scannevin RH, Huganir RL. 2000. Postsynaptic organisation and regulation of excitatory synapses. *Nature Reviews Neuroscience*, **1**(2): 133−141.

Schoch HJ, Fischer S, Marti HH. 2002. Hypoxia‐induced vascular endothelial growth factor expression causes vascular leakage in the brain. *Brain*, **125**(11): 2549−2557.

Semenza GL. 2009. Regulation of oxygen homeostasis by hypoxia-inducible factor 1. *Physiology*, **24**(2): 97−106.

Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB. 2010. Genetic evidence for high-altitude adaptation in Tibet. *Science*, **329**(5987): 72−75.

Singh M, Su C. 2013. Progesterone and neuroprotection. *Hormones and Behavior*, **63**(2): 284−290.

Soto C, Estrada LD. 2008. Protein misfolding and neurodegeneration. *Archives of Neurology & Psychiatry*, **65**(2): 184−189.

Sparkenbaugh E, Pawlinski R. 2013. Interplay between coagulation and vascular inflammation in sickle cell disease. *British Journal of Haematology*, **162**(1): 3−14.

Stamatovic SM, Dimitrijevic OB, Keep RF, Andjelkovic AV. 2006. Inflammation and brain edema: new insights into the role of chemokines and their receptors. *Acta Neurochirurgical Supplementum*, **96**: 444.

Stein DG. 2008. Progesterone exerts neuroprotective effects after brain injury. *Brain Research Reviews*, **57**(2): 386−397.

Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology*, **31**(1): 46−53.

Ward Jr JH. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58**(301): 236−244.

Wilson MH, Newman S, Imray CH. 2009. The cerebral effects of ascent to high altitudes. *The Lancet Neurology*, **8**(2): 175−191.

Won S, Sayeed I, Peterson BL, Wali B, Kahn JS, Stein DG. 2015. Vitamin D prevents hypoxia/reoxygenation-induced blood-brain barrier disruption via Vitamin D receptor-mediated NF-kB signaling pathways. *PLoS One*, **10**(3): e0122821.

Wright AD, Beazley MF, Bradwell AR, Chesner IM, Clayton RN, Forster PJ, Hillenbrand P, Imray CH, Society BMRE. 2004. Medroxyprogesterone at high altitude. *The effects on blood gases, cerebral regional oxygenation, and acute mountain sickness. Wilderness & Environmental Medicine*, **15**(1): 25−31.

Wroge CM, Hogins J, Eisenman L, Mennerick S. 2012. Synaptic NMDA receptors mediate hypoxic excitotoxic death. *Journal of Neuroscience*, **32**(19): 6732−6742.

Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, Van Gool AJ, Du H, Chen J, Zhang P, Huang Z, Thompson JR, Meng Y, Bai Y, Wang J, Zhuo M, Wang T, Huang Y, Wei L, Li J, Wang Z, Hu H, Yang P, Le L, Stenson PD, Li B, Liu X, Ball EV, An N, Huang Q, Zhang Y, Fan W, Zhang X, Li Y, Wang W, Katze MG, Su B, Nielsen R, Yang H, Wang J, Wang X, Wang J. 2011. Genome sequencing and comparison of two non-human primate animal models, the cynomolgus and Chinese rhesus macaques. *Nature Biotechnology*, **29**(11): 1019−1023.

Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, Li K, Luo R, Nie X, Wu H, Zhao M, Cao H, Zou J, Shan Y, Li S, Yang Q, As an, Ni P, Tian G, Xu J, Liu X, Jiang T, Wu R, Zhou G, Tang M, Qin J, Wang T, Feng S, Li G, Huasang, Luosang J, Wang W, Chen F, Wang Y, Zheng X, Li Z, Bianba Z, Yang G, Wang X, Tang S, Gao G, Chen Y, Luo Z, Gusang L, Cao Z, Zhang Q, Ouyang W, Ren X, Liang H, Zheng H, Huang Y, Li J, Bolund L, Kristiansen K, Li Y, Zhang Y, Zhang X, Li R, Li S, Yang H, Nielsen R, Wang J, Wang J. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, **329**(5987): 75−78.

Dysbiosis of gut microbiome affecting small intestine morphology and immune balance: a rhesus macaque model

Hong-Zhe Li[1,2,#], Nan Li[1,2,#], Jing-Jing Wang[1,2], Heng Li[1,2], Xing Huang[1,2], Lei Guo[1,2], Hui-Wen Zheng[1,2], Zhan-Long He[1], Yuan Zhao[1], Ze-Ning Yang[1,2], Hai-Tao Fan[1,2], Man-Man Chu[1,2], Jin-Xi Yang[1,2], Qiong-Wen Wu[1,2], Long-Ding Liu[1,2,*]

[1] *Institute of Medical Biology*, *Chinese Academy of Medical Sciences* & *Peking Union Medical College*, *Kunming*, *Yunnan* 650118, *China*
[2] *Key Laboratory of Systemic Innovative Research on Virus Vaccine*, *Chinese Academy of Medical Sciences*, *Kunming*, *Yunnan* 650118, *China*

ABSTRACT

There is a growing appreciation for the specific health benefits conferred by commensal microbiota on their hosts. Clinical microbiota analysis and animal studies in germ-free or antibiotic-treated mice have been crucial for improving our understanding of the role of the microbiome on the host mucosal surface; however, studies on the mechanisms involved in microbiome-host interactions remain limited to small animal models. Here, we demonstrated that rhesus monkeys under short-term broad-spectrum antibiotic treatment could be used as a model to study the gut mucosal host-microbiome niche and immune balance with steady health status. Results showed that the diversity and community structure of the gut commensal bacteria in rhesus monkeys were both disrupted after antibiotic treatment. Furthermore, the 16S rDNA amplicon sequencing results indicated that *Escherichia-Shigella* were predominant in stool samples 9 d of treatment, and the abundances of bacterial functional genes and predicted KEGG pathways were significantly changed. In addition to inducing aberrant morphology of small intestinal villi, the depletion of gut commensal bacteria led to increased proportions of CD3[+] T, CD4[+] T, and CD16[+] NK cells in peripheral blood mononuclear cells (PBMCs), but decreased numbers of Treg and CD20[+] B cells. The transcriptome of PBMCs from antibiotic-treated monkeys showed that the immune balance was affected by modulation of the expression of many functional genes, including IL-13, VCAM1, and LGR4.

Keywords: Gut microbiome; Rhesus macaque; Antibiotic treatment; Immune response; Pathological changes

INTRODUCTION

There is growing appreciation for the importance of commensal microbiota in shaping host development and physiology (Hooper & Gordon, 2001; Schmidt et al., 2018). Critically, the commensal microbiome is an important regulator of anti-infection immunity and colonizes the host for its lifetime (Abt et al., 2012; Postler & Ghosh, 2017). Several reports and

clinical cases have indicated an imbalance in immune cell subsets and abnormal up-regulation of some cytokines following gut microbiome dysbiosis, which is defined as an altered state of the microbial community (Ivanov & Honda, 2012; Langhorst et al., 2009; Soderborg & Friedman, 2018; Sprouse et al., 2019; Yang et al., 2015). These data are also supported by intestinal histopathological and immunological characteristics from experiments using mice treated with antibiotics (Kernbauer et al., 2014; Thackray et al., 2018), suggesting that the interaction between the microbiome and host shaped by commensal colonization could lead to a unique immune response. Studies on germ-free and antibiotic-treated mice have greatly improved our understanding of the specific health benefits conferred by commensal microbiota on the host immune system, digestive system, metabolic system, inflammation, and brain function (Belkaid & Hand, 2014; Desbonnet et al., 2015; Ferrer et al., 2014; Shapiro et al., 2014). However, in-depth studies on this interaction in mice have not provided sufficient data to help clarify the pathogenesis of immune disorders related to changes in the mucosal niche in the gut. Mechanisms that facilitate the establishment and stability of the gut microbiota remain poorly described. Moreover, rodent models have shortcomings such as high mortality, instability, autoimmune defects, and non-transformability, which limit the application of these research results (Vandamme, 2015). Thus, there is an urgent need for a suitable non-human primate (NHP) animal model to study the interactions among commensal microbiota and hosts.

Previous experimental results have indicated an essential role of the gut microbiome and probiotics in the intestinal mucosal barrier (Allaire et al., 2018; Yousefi et al., 2019). Furthermore, gene transcription profiling has been used to understand the systemic immune response that correlates with the function of microbiota in regulating the immune system (Gury-BenAri et al., 2016). Both innate immune and inflammatory responses after bacterial dysbiosis, which are responsible for regulating the variable integrated functions of the immune system, should be emphasized (Belkaid & Hand, 2014). In this work, the process of gut bacteria dysbiosis was verified using 12-month-old rhesus macaques (*Macaca mulatta*). Three rhesus monkeys were treated with a combination of antibiotics for 21 d to observe dynamic immunology and pathology process characteristics. Results showed that both the diversity and structure of the commensal bacteria were disrupted, and intestinal commensal bacterial species decreased by 95.6%–98.7% after antibiotic treatment. This depletion of gut commensal bacteria induced aberrant small intestinal morphology. In addition, we systemically analyzed and correlated the immune response and modulation of gene expression in peripheral blood mononuclear cells (PBMCs) from these monkeys. Gene transcripts in PBMCs that were up- and down-regulated after antibiotic treatment were identified, including those categorized under the GO terms of immune system process, cell communication, and cell activation. The results of this study may provide support for evaluating how changes in commensal bacteria affect

immune status, as reflected by PBMCs. Moreover, our study indicates that long-term oral antibiotic treatment could result in neurological symptoms. Therefore, the relationship between gut commensal bacteria and host immunity could be studied using this NHP model.

MATERIALS AND METHODS

Animals
Four one-year-old male rhesus monkeys (ID No.: 1, 2, 3, 4) were reared separately in a large cage (BSL-2 conditions) with sufficient fresh air and natural light, allowing visual, olfactory, and auditory interactions with other monkeys. All rhesus monkeys were healthy and weighed 2±0.5 kg. A temperature control valve was installed in each room to guarantee a room temperature of ~25 °C, with food, water, and fruit readily available. All animal experiments were performed with approval of the Yunnan Provincial Experimental Animal Management Association (Approval No.: SYXK (Dian) K2015-0006) and the Experimental Animal Administration and Ethics Committee of the Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College (Approval No.: DWSP201803006) and in accordance with the principles of the "Guide for the Care and Use of Laboratory Animals" and "Guidance to Experimental Animal Welfare and Ethical Treatment". All animals were fully under the care of veterinarians at the Institute of Medical Biology, Chinese Academy of Medicine Science.

Antibiotic treatment
Rhesus monkeys (ID No.: 1, 2, 3) were treated with a cocktail of antibiotics (1 g of ampicillin, 1 g of kanamycin, 1 g of metronidazole, 1 g of neomycin, and 1 g of vancomycin (Sigma-Aldrich, USA)) in 15 mL of 10% sucrose solution per day for three weeks, while ensuring health. The antibiotics were chosen based on previous research (Kernbauer et al., 2014) and taken orally, with care taken to avoid any confounding effects resulting from chronic stress caused by oral administration. The fourth rhesus monkey was treated with 10% sucrose solution per day for two weeks.

Sample collection and DNA extraction
Conventional animal samples were collected from other healthy rhesus monkeys of the same age. Fecal samples were collected every morning and frozen at −80 °C within 1 h of sampling. Blood samples were collected from the femoral or saphenous veins of monkeys without anesthesia every few days. Blood was divided into EDTA-K$_2$ tubes for cell detection and into separate serum tubes for the detection of cytokines. One of the antibiotic-treated monkeys (ID No. 2) was anesthetized with isoflurane, with blood then collected, followed by sacrifice via exsanguination. All tissues and contents were flash frozen in liquid nitrogen and stored at −80 °C until use, with a portion of the tissues fixed in formalin for histomorphometric analyses. Genomic DNA was extracted from stool using the QIAamp Fast DNA Stool Mini Kit (Qiagen, Germany) following standard protocols. The concentration of

genomic DNA was measured using a NanoDrop 2000 (Thermo Fisher Scientific, USA).

Hematology

A complete blood count (CBC) analysis was performed on whole blood immediately after sampling on day 14 after commencement of antibiotic treatment using an automatic hematological analyzer (XT-2000IV, Sysmex Corporation, Japan). Hematological values included white blood cells (WBCs), red blood cells (RBCs), neutrophils, lymphocytes, monocytes, eosinophils, basophils, mean corpuscular hemoglobin concentration (MCHC), and platelets.

16S rDNA amplicon sequencing and analyses

Polymerase chain reaction (PCR) amplification was performed to generate amplicons using bar-coded primers targeting the V3-V4 region of the bacterial 16S rRNA gene. The primers included 338F 5'-ACTCCTACGGGAGGCAGCA-3' and 806R 5'-GGACTACHVGGGTWTCTAAT-3'. Amplicons were extracted from 2% agarose gels, purified by the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, USA), and quantified using QuantiFluor™-ST (Promega Biosciences LLC, USA) according to standard protocols. The purified amplicons were then pooled and analyzed using a TruSeq™ DNA Sample Prep Kit and paired-end sequenced (2×300) on an Illumina MiSeq platform (Illumina Inc., USA) according to the manufacturer's instructions. Pairs of reads obtained by MiSeq sequencing were merged based on the overlapping relationship of paired-end reads. The quality of reads and the effect of merging were filtered by quality control. Effective sequences for each sample were obtained as operational taxonomic units (OTUs), which were clustered with a cut-off of 97% similarity using Usearch (v7.0, http://drive5.com/uparse), with single sequences removed. RDP Classifier (v2.2, http://sourceforge.net/projects/rdp-classifier/) was used to analyze the taxonomy of each representative OTU sequence against the SILVA (Release 128) 16S rRNA database with a confidence threshold of 0.7 (Amato et al., 2013; Quast et al., 2013). The community composition of each sample was counted at each taxonomic level. Alpha (α) diversity was analyzed by Mothur (Schloss et al., 2009), and statistical significance between two groups was calculated using Student's *t*- tests. Beta (β) diversity was estimated by computing the weighted UniFrac distance metric (Lozupone et al., 2006). Principal component analysis (PCA) was conducted according to Euclidean distance. Functional prediction analysis was performed to normalize the OTU abundance table by PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States, which stores Cluster of Ortholog Genes (COG) and KEGG Ortholog (KO) information corresponding to Greengene ID), removing the influence of the 16S marker gene on the number of copies in the genome, then obtaining COG family and KO information corresponding to each OTU and calculating the abundance.

Real-time PCR (RT-PCR) validation

Changes in gut bacterial community were validated by RT-

PCR using a SYBR Premix Ex Taq II kit (TaKaRa, Japan). We used DNA extracted from fecal samples as the template. Primers were designed to amplify specific regions in the hypervariable region V3 of the bacterial 16S rRNA gene (5'-ATTACCGCGGCTGCTGG-3' (F) and 5'-CTACGGAGGCAGC AG-3' (R)), *Enterobacteriaceae* (5'-CATTGACGTTACCCGC AGAAGAAGC-3' (F) and 5'-CTCTACGAGACTCAAGCTTGC-3' (R)), *Bacteroides-Prevotella* (5'-GAAGGTCCCCCACATTG-3' (F) and 5'-CAATCGGAGTTCTTCGTG-3' (R)), *Bifidobacterium* (5'-GGGTGGTAATGCCGGATG- 3' (F) and 5'-TAAGCCATGGACTTTCACACC-3' (R)), *Lactobacillus* (5'-AGC AGTAGGGAATCTTCCA-3' (F) and 5'-ATTYCACCGCTACAC ATG-3' (R)), *Enterococcus-Vagococcus* (5'-AACCTACCC ATCAGAGGG-3' (F) and 5'-GACGTTCAGTTACTAACG-3' (R)), *Clostridium phoceensis* (5'-GATGGCCTCGCGTCCG ATTAG-3' (F) and 5'-CCGAAGACCTTCTTCCTCC-3' (R)), *Clostridium cluster I* (5'-TACCHRAGGAGGAAGCCAC- 3' (F) and 5'-GTTCTTCCTAATCTCTACGCAT-3' (R)), *Clostridium cluster XI* (5'-ACGCTACTTGAGGAGGA- 3' (F) and 5'-GAGCCGTAGCCTTTCACT-3' (R)), and *Clostridium cluster XIV* (5'-GAWGAAGTATYTCGGTATGT- 3' (F) and 5'-CTACGCWCCCTTTACAC-3' (R)). We used the $2^{-\Delta Ct}$ method to calculate the richness of the gut bacteria.

Flow cytometry and LiquiChip

Peripheral blood (100 μL) was incubated with antibodies for 30 min at room temperature in the dark and then incubated with red blood cell lysis buffer and washed with phosphate-buffered saline (PBS). The cells were analyzed using a CytoFLEX flow cytometer (Beckman Coulter, USA) according to the manufacturer's instructions. The following antibodies (clones) were used for staining: CD3 (SP34-2), CD4 (L200), CD20 (2H7), CD8 (RPA-T8), and CD25 (M-A251), all from BD Bioscience (USA). All antibodies were titered in advance and used at optimal concentrations for flow cytometry. FlowJo v.10 was used to analyze the data. Measurement of cytokines and chemokines in the serum was performed using a MILLIPLEX® MAP NHP cytokine magnetic bead panel kit (Millipore Corporation, USA) and detected by a Bio-Plex 200 System (Bio-Rad Laboratories, USA).

Agilent genome microarray

The PBMCs were isolated by density gradient centrifugation with Lymphoprep medium (Ficoll-Paque PREMIUM; GE Healthcare, USA). Total RNA was extracted using TRIzol Reagent (Cat#15596-018, Life Technologies, USA), following the manufacturer's instructions, and checked for RNA integrity numbers (RIN) to determine integrity using an Agilent Bioanalyzer 2100 (Agilent Technologies, USA). Microarray analysis was performed using the Agilent Rhesus Macaque Genome Microarray (USA, 4×44K). The arrays were hybridized, washed, and scanned according to the standard protocols. Gene chip tests were performed by the Shanghai Biochip Company (China). Data were extracted with Feature Extraction software v10.7 (Agilent Technologies, USA). Raw data were normalized by the quantile algorithm limma packages in R. The log-transformed expression values were

adjusted, and fold-change statistical method was used to select differentially expressed genes (DEGs). Gene Ontology (GO), pathway enrichment, and network analysis of significant DEGs were systematically conducted.

Morphology
The small intestine and brain were fixed in formalin and embedded in paraffin according to standard histological protocols. Paraffin-embedded sections were deparaffinized and stained with hematoxylin-eosin-safran (H&E). The slides were scanned by a Pannoramic MIDI scanner (3DHISTECH, Hungary), and all measurements were made using CaseViewer software v2.2.

Data analysis
Significant differences between the values of two groups were calculated using Student's *t*- tests. IBM SPSS statistics v23 was used for data evaluation. Statistical significance was considered at $P<0.05$.

RESULTS

Health status of rhesus monkeys during oral antibiotic treatment
The three rhesus monkeys exhibited normal food and water intake, activities, and mental state during antibiotic treatment. Diarrhea symptoms appeared on day 2 after antibiotic treatment, which recovered after one week (Table 1). Hematological levels remained normal after 14 d of antibiotic (ABX) treatment (Figure 1). Compared with conventional (Conv) rhesus monkeys, the mean values for white blood cells, neutrophils, lymphocytes, and mean corpuscular hemoglobin concentration (MCHC) were significantly lower in rhesus monkeys treated with antibiotics (Figure 1A, B). Therefore, the body status tended to be stable after 14 d of antibiotic treatment.

Table 1 Antibiotic treatment induces diarrhea symptoms in rhesus monkeys

Days post-ABX treatment	Severity of diarrhea		
	1	2	3
0	−	−	−
1	++	++	++
2	+++	+++	+++
3	+++	+++	+++
4	+++	+++	+++
5	+++	+++	+++
6	++	++	++
7	+	+	+
8	−	+	−
9	−	−	−
10	−	−	−

1, 2, 3 are IDs of rhesus monkeys. Severity of diarrhea was determined by number of defecations, characteristics of feces, and degree of dehydration. +: Mild diarrhea; ++: Diarrhea; +++: Severe diarrhea; −: Normal stool.

Antibiotic-treated rhesus monkeys display a dramatic shift in bacterial community structure in the intestine
Antibiotic treatment resulted in a significant reduction in the abundance of commensal bacteria and reorganization of bacterial composition. According to the results of 16S rDNA amplicon sequencing, 538 722 effective sequences were obtained from samples, with an average length of 448 nt. Taxonomic analysis of OTUs resulted in 1, 1, 12, 20, 37, 61, 143, 218, and 324 different categories at the domain, kingdom, phylum, class, order, family, genus, species, and OTU levels, respectively. Based on α diversity analysis, the abundance and diversity of each sample were sufficient to fully describe the composition of the bacteria, and the sequencing depth was higher than 0.99 (Table 2). After antibiotic treatment, the Shannon index decreased significantly, and the number of OTUs decreased by 96.0%−98.6% after 21 d (Table 2). At the phylum level, the intestinal commensal bacteria in untreated rhesus monkeys mainly consisted of *Bacteroidetes* and *Firmicutes* (Figure 2A), and dominant bacteria included *Prevotella*, *Lactobacillus*, *Bacteroides*, *Lachnospira*, *Phascolarctobacterium*, *Faecalibacterium*, *Ruminococcus*, *Megasphaera* , and *Subdoligranulum* (Figure 2B). The diversity and abundance of the commensal bacteria decreased significantly after 21 d; only a few *Proteobacteria* were detected, and most were antibiotic-resistant *Escherichia* and *Shigella* (Figure 2B, C). Based on statistical analysis and PCA, we identified a significant difference in the structure of the commensal bacterial community between untreated and antibiotic-treated rhesus monkeys, and the reorganized bacterial structure maintained stability (Figure 2D, E). We next investigated the enterotype according to the clustering of dominant bacterial communities (Arumugam et al., 2011). Results showed that the bacterial communities were most naturally categorized into eight clusters during treatment (Figure 2F), and antibiotic treatment in rhesus monkeys resulted in a change from *Prevotella* enterotype to *Escherichia-Shigella* enterotype (Figure 2G). In addition, the diversity and community structure of the intestinal bacteria in the monkey (ID No. 4) treated only with sucrose remained unchanged (Figure 2H).

To verify the 16S rDNA amplicon sequencing results, we used quantitative RT-PCR to detect the composition and abundance of intestinal bacteria in the three rhesus monkeys. Results showed that the dominant commensal bacteria in untreated rhesus monkeys consisted of *Bacteroides*, *Prevotella*, *Lactobacillu* s, and *Clostridium* (Supplementary Figure S1A–C). We also tested antibiotic resistance by selective agar medium, and found no antibiotic-resistant strain in the normal commensal bacteria. After 3 d of antibiotic treatment, the copies of commensal bacteria per milligram of feces decreased by 10 000 to 100 000 times (Supplementary Figure S2). After 14 d of antibiotic treatment, the abundance of the 16S rRNA gene decreased significantly in the feces, and predominant bacteria were depleted (Supplementary Figure S1D –F). Furthermore, antibiotics induced the rapid rise of resistant *Escherichia coli* and *Shigella,* which did not exist
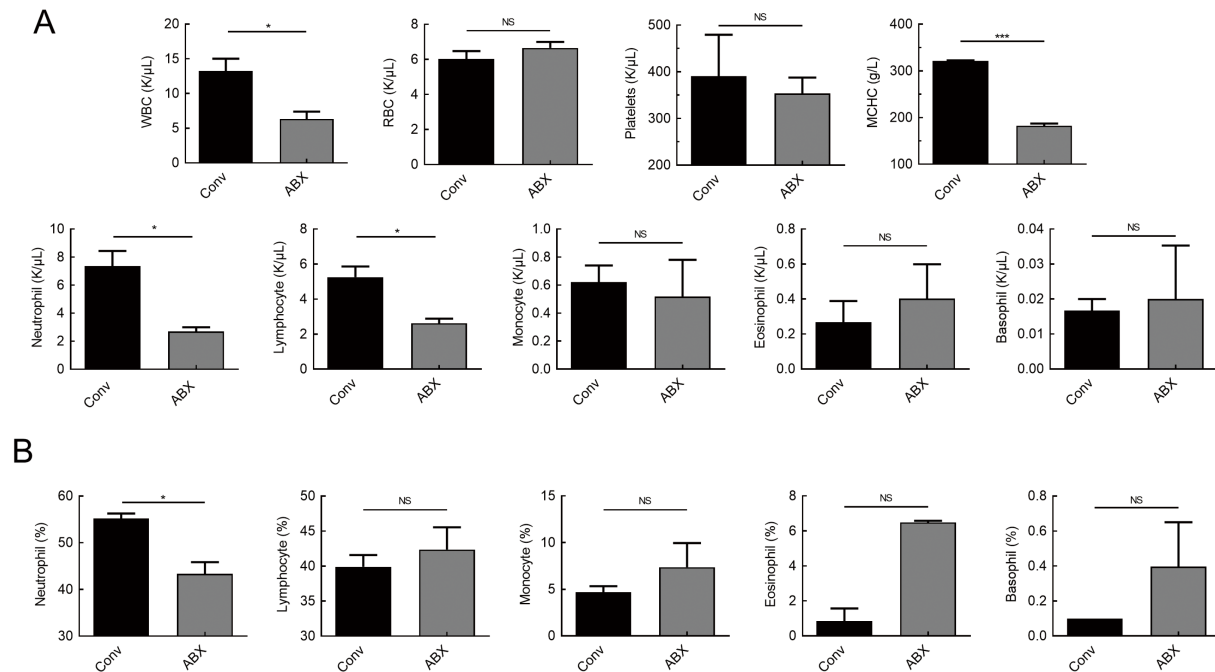
Figure 1 Hematological changes in rhesus monkeys after 14 d of antibiotic treatment

A: Hematological values of white blood cells (WBC), red blood cells (RBC), platelets, mean corpuscular hemoglobin concentration (MCHC), neutrophils, lymphocytes, monocytes, eosinophils, and basophils. B: Proportion of neutrophils, lymphocytes, monocytes, eosinophils, and basophils in blood. Conv: Conventional or untreated rhesus monkeys. Graphs show means±standard error of mean (*SEM*). *n*=3 rhesus monkeys per group. Significant differences between values were calculated using an unpaired two-tailed *t*-test. NS: Not significant, *: *P*<0.05, **: *P*<0.01, ***: *P*<0.001.

Table 2 Alpha diversity estimators of 16S rDNA amplicon sequencing

ID_days post ABX treatment	Abundance index			Diversity index		Coverage
	Sobs	Ace	Chao	Shannon	Simpson	
1_D0	284	286.31	286.55	4.175 399	0.045 147	0.999 685
1_D5	28	86.14	41	0.101 25	0.966 94	0.999 573
1_D14	13	18.79	16.33	0.011 107	0.997 727	0.999 874
1_D21	4	5.59	4	0.003 219	0.999 342	0.999 97
2_D0	227	247.43	254	3.398 891	0.112 415	0.999 196
2_D5	8	43.56	11	0.477 874	0.702 516	0.999 877
2_D14	7	15.84	8.5	0.005 4	0.998 908	0.999 918
2_D21	6	6.83	6	0.009 683	0.997 853	0.999 977
3_D0	249	256.00	260.33	3.843 893	0.060 07	0.999 246
3_D5	28	117.74	49	0.190 655	0.926 19	0.999 52
3_D14	32	149.41	95.33	0.027 369	0.994 332	0.999 415
3_D21	10	91.57	17.5	0.013 941	0.996 685	0.999 844

Data were divided as individuals described in the text. Number of alpha diversity indices are shown.

before (Supplementary Figure S1G–I).

Predictive functional profiling changes driven by microbial shifts

After standardization of OTU abundances, the 16S functional predictions were obtained from COG family information according to the corresponding Greengene ID of each OTU. Functional abundances were acquired by analyzing the descriptive and functional information of COG classifications in the eggNOG database. According to the analysis of bacterial functional genomics, we found that the abundance of functional genes in RNA processing and modification, cell motility, and intracellular trafficking were the most significantly increased (*P*<0.001) after antibiotic treatment (Figure 3A).

To characterize the functional alterations of commensal bacteria in antibiotic-treated rhesus monkeys, we predicted the functional composition profiles from the 16S rDNA

Figure 2 Changes in commensal bacteria spectrum in stool after antibiotic treatment

A: Relationship between samples and bacterial community at phylum level. Data were rendered by Circos (Krzywinski et al., 2009). Left half-circle represents composition of species in sample, color of outer ribbon represents which group it comes from, and colors of inner ribbon represent phyla. Right half-circle represents distribution proportion of each phylum in samples from different days, outer ribbons represent phyla, inner ribbon color represents different groups. Length of bars from each phylum indicates relative abundance of that phylum in corresponding sample. B: Abundances of top 35 genera in each sample were selected and compared with abundances of these genera in other samples by heatmap. C: Network analysis elucidating distribution of samples and species (abundance>50) at OTU level, highlighting similarities and differences between samples. D: PCA plot displaying variation in community structure during treatment at OTU level. Each point represents an individual. E: Bar plot showing significantly different phylotypes between pre- and post-ABX-treated rhesus monkeys at genus level. Statistical analysis was performed by Student's *t*-test and *P*-values were adjusted by FDR. *n* =3, in each group. *: *P* <0.05, **: *P* <0.01, ***: *P* <0.001. F, G: Enterotype analysis. Clustering analysis using Jensen-Shannon distance (JSD) and partitioning around medoids (PAM) method. Calinski-Harabasz (CH) index was then used to calculate optimal clustering *K* value (*K*=8), and bar plot was used for visualization of each sample's enterotype. H: Bar plot displaying variation in community structure of monkey receiving sucrose-treatment only at genus level.

sequencing data under PICRUSt pre- and post-antibiotic treatment. We found that multiple KEGG (level 2) categories were disturbed. The changes in pathways enriched in signal transduction, excretory system, neurodegenerative diseases, infectious diseases, and cell motility showed the most significant differences (*P*<0.001) between untreated and antibiotic-treated rhesus monkeys (Figure 3B). Strikingly, abundances in neurodegenerative disease, infectious disease, cancer, and metabolism pathways were significantly increased after commensal bacteria depletion. In addition, abundances in the digestive system pathway decreased along with signaling molecules and interaction pathways.

Impact of commensal bacteria alteration on host immune profile in PBMCs

As the immune system is modulated by the gut microbiota, we examined whether oral antibiotic treatment affected lymphocyte populations in peripheral blood. We observed increased numbers of CD3[+] T cells and CD16[+] NK cells after antibiotic treatment, as well as greater CD4[+] and CD8[+] cells (Figure 4B). In contrast to CD3[+] T cells, we observed decreased numbers of Treg cells and CD20[+] B cells after antibiotic treatment, suggesting a potential defect in the humoral immune response (Figure 4B). In the cytokine expression profile, we found that CD40L maintained a stable level in serum, indicating a pivotal role in co-stimulation and regulation of the adaptive immune response. The concentration of inflammatory cytokines also varied with the development of intestinal bacterial depletion (Figure 4C).

In addition, we performed an Agilent genome microarray with PBMC samples to acquire gene expression profiling of immune cells. In biological process analysis, we focused on the metabolic process, immune system process, cell communication, and cell activation GO terms. The GO terms of leukocyte activation, immune response, cell-cell signaling, and thyroid hormone metabolic process were significantly different pre- and post-antibiotic treatment (Figure 5). In immune system response and cell-cell signaling, the expression level of LGR4 decreased more than 10-fold. The protein encoded by this gene is a G-protein coulped receptor

that binds R-spondins, activates the Wnt signaling pathway, and is associated with osteoporosis (Styrkarsdottir et al., 2013). The IFNA2, CXCL10, and TLR7 genes were also down-regulated, which may affect host susceptibility to viral infection (Karst, 2016; Spurrell et al., 2005; Wu et al., 2013). At the same time, the levels of VCAM1 and IL-13 were upregulated more than 5-fold (Figure 5A, C). It has been reported that IL-13 induces several changes in the gut that can lead to detachment of organisms from the gut wall (Seyfizadeh et al., 2015), and overexpression of IL-13 may contribute to some features of allergic lung diseases such as airway hyperresponsiveness, goblet cell metaplasia, and mucus hypersecretion (Wills-Karp et al., 1998). Several effector factors of immune checkpoint proteins, tissue remodeling, and cell adhesion, such as the MMP14, ABL1, and LILRA6 genes, were markedly elevated. Additionally, several cell communication genes, including GRIA2, BCAN, CACNG3, SLC12A5, SOX17, TDGF1, and VCAM1, were up-regulated (Figure 5C). These results suggest that the absence of commensal bacteria may alter the immune system and disease status of the host.

Depletion of commensal bacteria impairs development of small intestinal morphology

Compared with the conventional rhesus macaques, the antibiotic-treated monkeys showed impaired morphology of the small intestine. The top of the villi of the intestinal mucosa were diminished, and the epithelial cells of the intestinal mucosa were denatured, necrotic, and shedding, forming extensive superficial erosion (Figure 6). Thus, our results demonstrated that dysbiosis of commensal bacteria could impair the small intestine, as reported in previous animal studies (Kernbauer et al., 2014; Yeruva et al., 2016).

Long-term use of antibiotics leads to severe adverse reactions in a rhesus monkey

Although drugs are a known cause of neurological symptoms, antibiotics have not been taken seriously and the frequency of severe central nervous system (CNS) events associated with antibiotics is reported to be less than 1% (Bhattacharyya et al.,
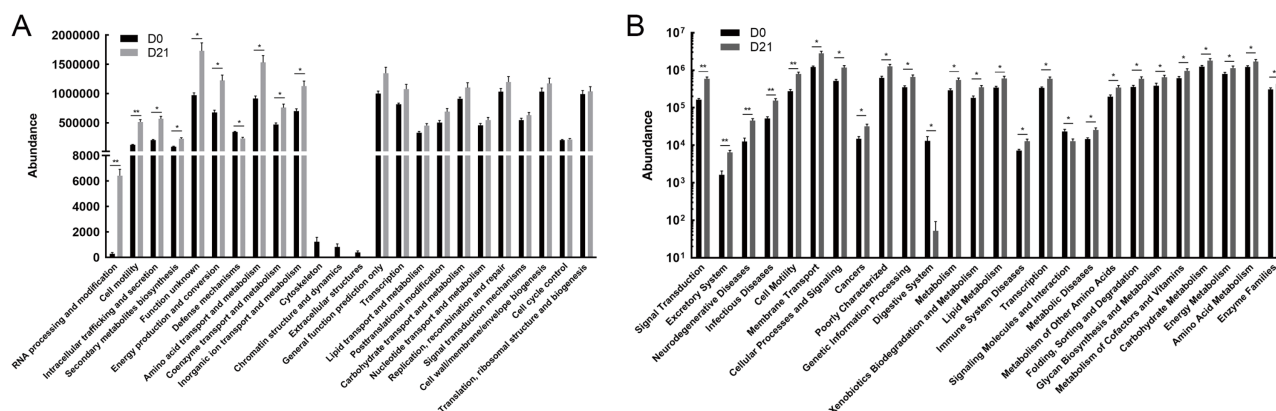

Figure 3 Microbial functions altered by restructured bacteria
COG function (A) and KEGG pathway (B) prediction analyses of bacterial microbiome. *n*=3 per group. *: *P*<0.05, **: *P*<0.01, ***: *P*<0.001.
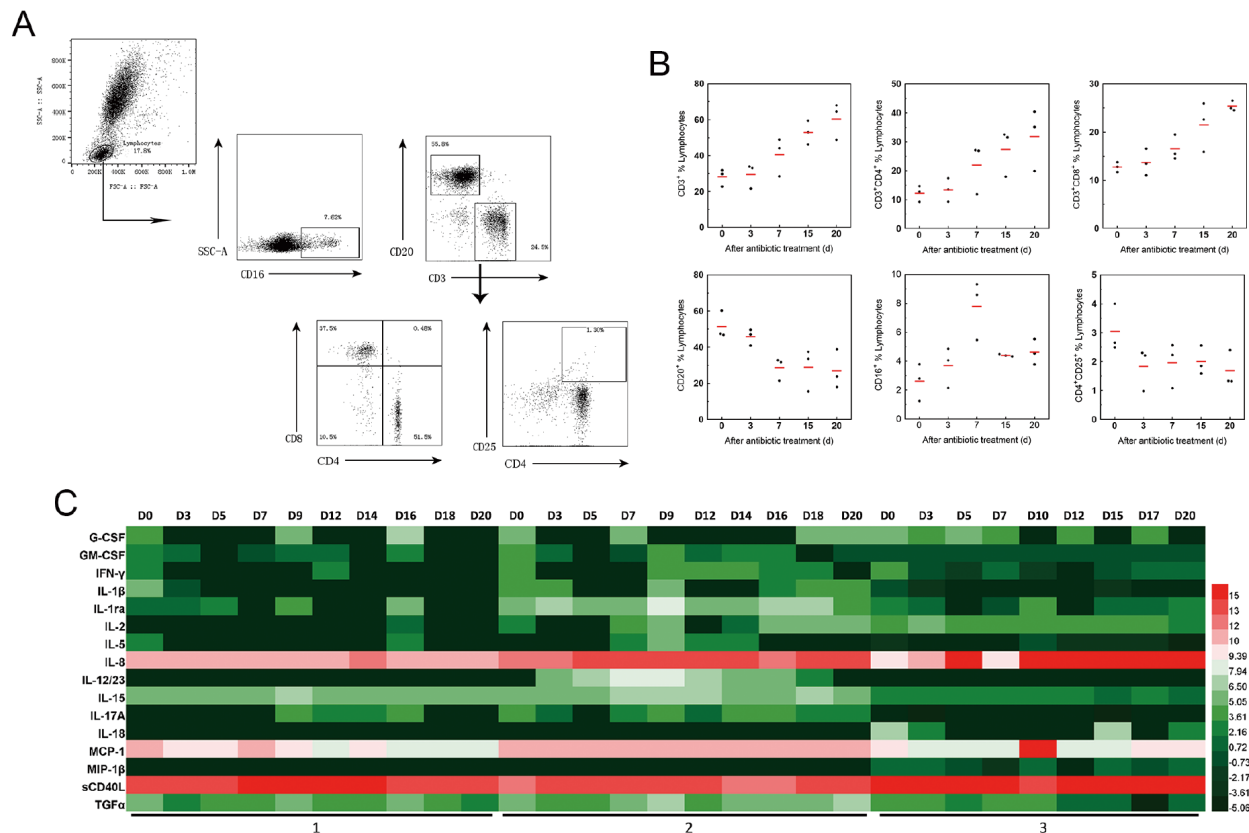
Figure 4 Host immune responses evoked by alterations in microbiome

A: Gating strategies for detection of lymphocyte subsets by flow cytometry. B: Changes in frequency of CD3$^+$, CD3$^+$CD4$^+$, CD3$^+$CD8$^+$, CD20$^+$, CD16$^+$, and CD4$^+$CD25$^+$ T cells among lymphocytes in PBMCs after antibiotic treatment. Red line represents mean value at each point. C: Heatmap displaying change in different cytokines in serum during antibiotic treatment. Concentrations (pg/mL) were detected by multiplex immunoassay and values were log2 transformed.

2016; Owens & Ambrose, 2005). However, the results of a recent retrospective study suggest that bacteria-associated neurogenic disease may be underestimated (Sandler et al., 2000). In our study, one rhesus monkey (ID No. 2) exhibited sudden myoclonus after 20 d of oral antibiotic treatment, which was improved by the reduction of the antibiotic to one third of the designed dose, although convulsive seizure occurred as long as antibiotic treatment continued. Finally, we found abnormal brain morphology in this rhesus monkey. The thalamus appeared congested (Figure 7B), and the area between the inner and outer bundles was sparse, similar to softening foci (Figure 7C).

DISCUSSION

Under conventional conditions, gut commensal bacteria inhabit the surface of the intestinal epithelium, forming a stable symbiotic relationship with the host and shaping a physical barrier on the surface (Baumgart & Dignass, 2002; Davenport et al., 2017). The presence of intestinal commensal bacteria reduces the colonization, translocation, and growth of

conditional pathogens (Kamada et al., 2013) and is therefore referred to as "colonization resistance". Dysbiosis or lack of commensal bacteria can cause extensive proliferation of pathogenic bacteria or other pathogens in the intestine, even causing diseases in the host (Thackray et al., 2018; Van den Bergh et al., 2016; Wienhold et al., 2018). With the development of genome sequencing technology, we are aware of the relationship between commensal bacteria and the host. However, the mechanism of interactions among dominant bacterial species, pathogens, and hosts remains unclear.

NHPs exhibit high similarity with humans in anatomical structure, physiological metabolism, and immune system. However, few studies have systematically compared gut commensal microbiomes between human and NHPs (Davenport et al., 2017). Rhesus macaques can still be used in basic studies that cannot be completed or simulated with rodent models. However, due to limitations of body size, feeding conditions, and growth conditions, no studies have reported on commensal microbiota using the rhesus monkey as an animal model. In addition to germ-free animal models, antibiotic-treated animal models are also commonly used to
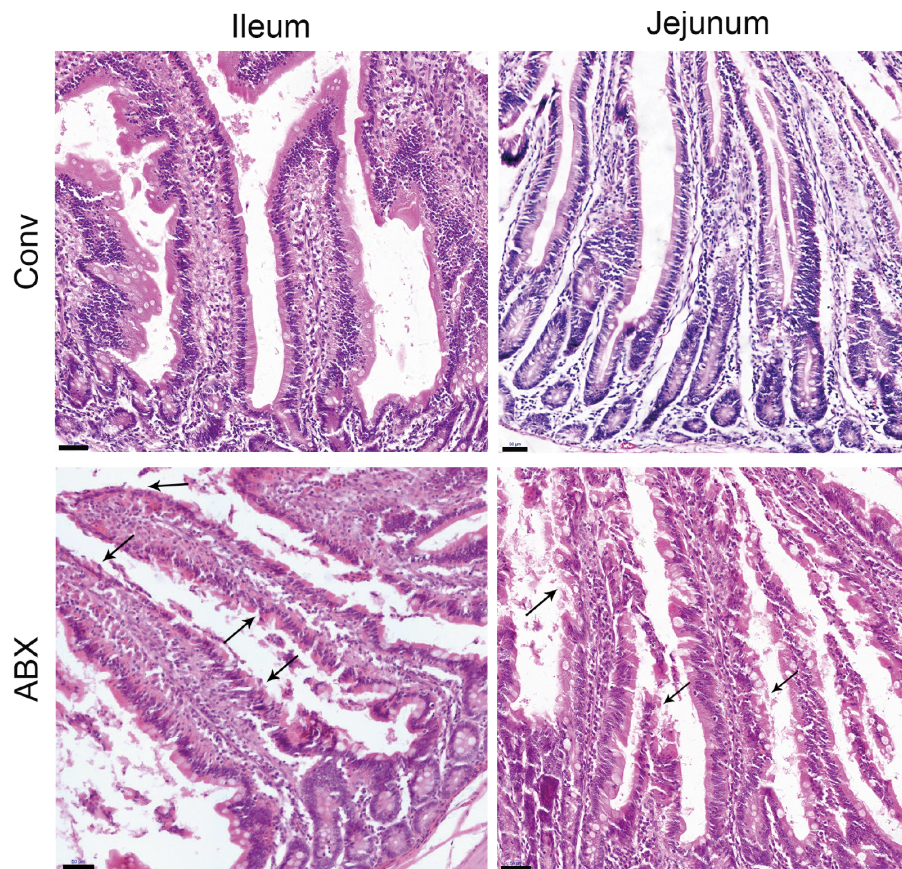
Figure 5 Network plots showing host-gene responses significantly regulated after antibiotic treatment

Co-occurrence analysis of significantly different GO terms and gene expression profiling, including immune system process (A), metabolic process (B), and cell communication (C). Square nodes represent GO terms, elliptical nodes represent functional genes, and different colors show different fold-change values.

study intestinal bacteria (Karst, 2016). In this study, four broad-spectrum antibiotics, including aminoglycoside, penicillin, glycopeptide, and nitroimidazole, were co-administered, and were found to deplete commensal bacteria in the intestines of rhesus monkeys. This study has some limitations, including that we were unable to completely deplete the intestinal bacterial community. The relative abundance of the 16S rRNA gene decreased by more than 80% on day 9 of treatment according to RT-PCR analysis, and the subsequent bacterial structure remained relatively stable. Based on the 16S rDNA amplicon sequencing results, the OTU level of intestinal symbionts decreased by more than 90% on day 9 after antibiotic treatment. After 14 d, the dominant bacterial group changed to *Escherichia-Shigella*, and both the richness and diversity of the intestinal bacterial community were depleted stably and continuously. Overall, the aim of eliminating enteric commensal bacteria in rhesus monkeys can be achieved 9 d after high-dose oral antibiotic treatment.

The pathological findings of the small intestine indicated that the absence of commensal bacteria can lead to aberrant intestinal morphology, including that of intestinal epithelial cells and villi. Flow cytometry results also showed alterations in lymphocyte populations, with increased numbers of T lymphocytes in peripheral blood. At the same time, the abundance of infectious diseases in KEGG prediction analysis of the microbiome significantly increased after antibiotic treatment. Thus, the lack of commensal bacteria could disrupt colonization resistance and affect the host immune system. Microarray analysis of the transcriptome performed on PBMCs isolated from antibiotic-treated monkeys also showed that the immune balance was affected by modulation in the expression of many different genes. The innate immunity-associated genes demonstrated cross-talk in response to bacterial dysbiosis with a high level of variation. These genes, including IL-13, CD28, CCR2, IL-12β, IL-1RL2, VCAM1, and FCER1A, exhibited obvious up-regulation with activation of the MHC system, T cell activity, and inflammatory responses. In addition, genes encoding cytokines and chemokines, such as IFNα2, CXCL10, and IL-4, were down-regulated. The alteration in the expression of these genes likely contributes to the effective functional activation of the innate immune response (Cross et al., 2004). These results suggest that bacterial dysbiosis caused by antibiotic treatment may affect the outcome of intestinal infectious diseases, including both bacterial and viral infections.

Furthermore, frequent oral antibiotic treatment could lead to the production of highly resistant *E. coli* (Van den Bergh et al., 2016). This study found that drug-resistant *E. coli* colonies appeared 4 d after antibiotic treatment, whereas no *Escherichia-Shigella* strains were present in the gut before treatment. This result should serve as an alarm suggesting that extensive clinical trials and optimization of antibiotic

Figure 6 Histopathological changes in small intestine in antibiotic-treated rhesus monkeys

Comparison of ileum and jejunum between conventional (Conv) and antibiotic-treated (ABX) rhesus monkeys. n =1 per group. Tissue sections stained with H&E. Images were taken at 20× magnification; Scale bars: 50 µm. Black arrows show aberrant intestinal morphology, including diminished top of villi and denatured epithelial cells of intestinal mucosa.
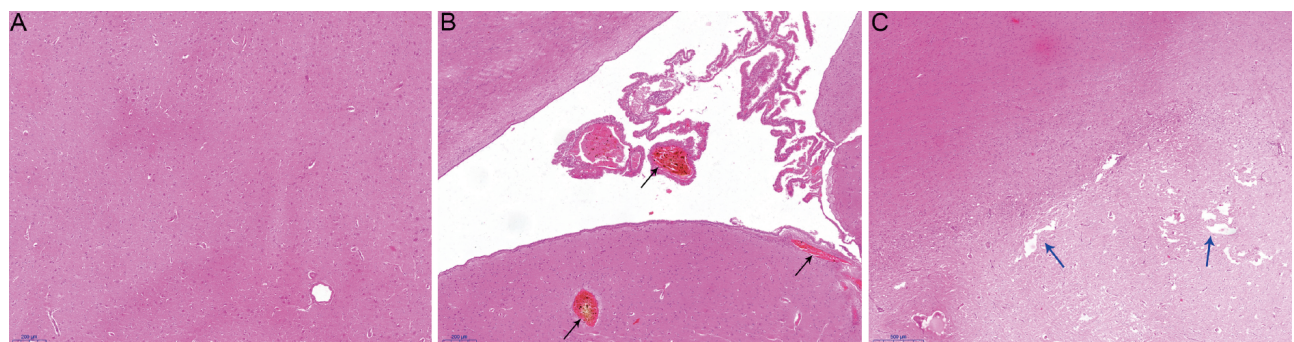

Figure 7 Histopathological changes in brain after long-term antibiotic treatment

A: Thalamus of conventional rhesus monkey. B, C: Congestion (black arrows) and lacunar state (blue arrows) pathological changes in thalamus of antibiotic-treated rhesus monkey. Images were taken at 20× magnification; Scale bars: 200 µm.

treatment should be conducted to minimize the possibility of strain resistance. In this study, the use of antibiotics caused adverse reactions such as diarrhea and vomiting in the short term. In addition, long-term antibiotic use resulted in serious symptoms such as myoclonus. Predictive functional profiling of bacteria also showed that abundance of the KEGG pathway of neurodegenerative diseases increased significantly, suggesting that the diversity and stability of intestinal commensal bacteria are important for the maintenance of the nervous system.

In conclusion, our data suggest that early childhood represent a critical period during which perturbations to the gut microbiota and dysregulation of microbiota-gut-immunity axis communication may significantly impact the host immune system in adulthood. Although the precise mechanisms by which gut microbiota mediate changes in the host have not yet been elucidated, our findings indicate that rhesus monkeys with short-term high-dose antibiotic treatment represent a useful model for assessing the importance of gut microbiota to the host during distinct stages of early life, without disruption to host health. This is an unavailable feature of the germ-free animal model. Further studies focusing on the effects of bacterial depletion on the host are required to identify its potential relevance to neurodevelopmental and infectious diseases and functional mechanisms.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

ACKNOWLEDGEMENTS

REFERENCES

Abt MC, Osborne LC, Monticelli LA, Doering TA, Alenghat T, Sonnenberg GF, Paley MA, Antenus M, Williams KL, Erikson J, Wherry EJ, Artis D. 2012. Commensal bacteria calibrate the activation threshold of innate antiviral immunity. *Immunity*, **37**(1): 158−170.

Allaire JM, Crowley SM, Law HT, Chang SY, Ko HJ, Vallance BA. 2018. The intestinal epithelium: Central coordinator of mucosal immunity. *Trends in Immunology*, **39**(9): 677−696.

Amato KR, Yeoman CJ, Kent A, Righini N, Carbonero F, Estrada A, Gaskins HR, Stumpf RM, Yildirim S, Torralba M, Gillis M, Wilson BA, Nelson KE, White BA, Leigh SR. 2013. Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes. *The ISME Journal*, **7**(7): 1344−1353.

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T,

Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Doré J, Meta HITC, Antolín M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C, Cultrone A, Delorme C, Denariaz G, Dervyn R, Foerstner KU, Friss C, van de Guchte M, Guedon E, Haimet F, Huber W, van Hylckama-Vlieg J, Jamet A, Juste C, Kaci G, Knol J, Lakhdari O, Layec S, Le Roux K, Maguin E, Mérieux A, Melo Minardi R, M'Rini C, Muller J, Oozeer R, Parkhill J, Renault P, Rescigno M, Sanchez N, Sunagawa S, Torrejon A, Turner K, Vandemeulebrouck G, Varela E, Winogradsky Y, Zeller G, Weissenbach J, Ehrlich SD, Bork P. 2011. Enterotypes of the human gut microbiome. *Nature*, **473**(7346): 174−180.

Baumgart DC, Dignass AU. 2002. Intestinal barrier function. *Current Opinion in Clinical Nutrition and Metabolic Care*, **5**(6): 685−694.

Belkaid Y, Hand TW. 2014. Role of the microbiota in immunity and inflammation. *Cell*, **157**(1): 121−141.

Bhattacharyya S, Darby RR, Raibagkar P, Gonzalez Castro LN, Berkowitz AL. 2016. Antibiotic-associated encephalopathy. *Neurology*, **86**(10): 963−971.

Cross ML, Ganner A, Teilab D, Fray LM. 2004. Patterns of cytokine induction by gram-positive and gram-negative probiotic bacteria. *FEMS Immunology & Medical Microbiology*, **42**(2): 173−180.

Davenport ER, Sanders JG, Song SJ, Amato KR, Clark AG, Knight R. 2017. The human microbiome in evolution. *BMC Biology*, **15**(1): 127.

Desbonnet L, Clarke G, Traplin A, O'Sullivan O, Crispie F, Moloney RD, Cotter PD, Dinan TG, Cryan JF. 2015. Gut microbiota depletion from early adolescence in mice: Implications for brain and behaviour. *Brain, Behavior, and Immunity*, **48**: 165−173.

Ferrer M, Martins dos Santos VA, Ott SJ, Moya A. 2014. Gut microbiota disturbance during antibiotic therapy: A multi-omic approach. *Gut Microbes*, **5**(1): 64−70.

Gury-BenAri M, Thaiss CA, Serafini N, Winter DR, Giladi A, Lara-Astiaso D, Levy M, Salame TM, Weiner A, David E, Shapiro H, Dori-Bachash M, Pevsner-Fischer M, Lorenzo-Vivas E, Keren-Shaul H, Paul F, Harmelin A, Eberl G, Itzkovitz S, Tanay A, Di Santo JP, Elinav E, Amit I. 2016. The spectrum and regulatory landscape of intestinal innate lymphoid cells are shaped by the microbiome. *Cell*, **166**(5): 1231−1246.

Hooper LV, Gordon JI. 2001. Commensal host-bacterial relationships in the gut. *Science*, **292**(5519): 1115−1118.

Ivanov II, Honda K. 2012. Intestinal commensal microbes as immune modulators. *Cell Host and Microbe*, **12**(4): 496−508.

Kamada N, Chen GY, Inohara N, Núñez G. 2013. Control of pathogens and pathobionts by the gut microbiota. *Nature Immunology*, **14**(7): 685−690.

Karst SM. 2016. The influence of commensal bacteria on infection with enteric viruses. *Nature Reviews Microbiology*, **14**(4): 197−204.

Kernbauer E, Ding Y, Cadwell K. 2014. An enteric virus can replace the beneficial function of commensal bacteria. *Nature*, **516**(7529): 94−98.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: An information aesthetic for comparative genomics. *Genome Research*, **19**(9): 1639−1645.

Langhorst J, Junge A, Rueffer A, Wehkamp J, Foell D, Michalsen A, Musial F, Dobos GJ. 2009. Elevated human beta-defensin-2 levels indicate an activation of the innate immune system in patients with irritable bowel syndrome. *The American Journal of Gastroenterology*, **104**(2): 404−410.

Lozupone C, Hamady M, Knight R. 2006. Unifrac--an online tool for

comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics*, **7**(1): 371.

Owens RC, Jr., Ambrose PG. 2005. Antimicrobial safety: Focus on fluoroquinolones. *Clinical Infectious Diseases*, **41**(Suppl 2): S144−157.

Postler TS, Ghosh S. 2017. Understanding the holobiont: How microbial metabolites affect human health and shape the immune system. *Cell Metabolism*, **26**(1): 110−130.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, **41**(D1): D590−D596.

Sandler RH, Finegold SM, Bolte ER, Buchanan CP, Maxwell AP, Väisänen ML, Nelson MN, Wexler HM. 2000. Short-term benefit from oral vancomycin treatment of regressive-onset autism. *Journal of Child Neurology*, **15**(7): 429−435.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**(23): 7537−7541.

Schmidt TSB, Raes J, Bork P. 2018. The human gut microbiome: From association to modulation. *Cell*, **172**(6): 1198−1215.

Seyfizadeh N, Seyfizadeh N, Gharibi T, Babaloo Z. 2015. Interleukin-13 as an important cytokine: A review on its roles in some human diseases. *Acta Microbiologica et Immunologica Hungarica*, **62**(4): 341−378.

Shapiro H, Thaiss CA, Levy M, Elinav E. 2014. The cross talk between microbiota and the immune system: Metabolites take center stage. *Current Opinion in Immunology*, **30**: 54−62.

Soderborg TK, Friedman JE. 2018. Imbalance in gut microbes from babies born to obese mothers increases gut permeability and myeloid cell adaptations that provoke obesity and NAFLD. *Microbial Cell*, **6**(1): 102−104.

Sprouse ML, Bates NA, Felix KM, Wu HJ. 2019. Impact of gut microbiota on gut-distal autoimmunity: A focus on t cells. *Immunology*, **156**(4): 305−318.

Spurrell JC, Wiehler S, Zaheer RS, Sanders SP, Proud D. 2005. Human airway epithelial cells produce ip-10(CXCL10) in vitro and in vivo upon rhinovirus infection. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, **289**(1): L85−L95.

Styrkarsdottir U, Thorleifsson G, Sulem P, Gudbjartsson DF, Sigurdsson A, Jonasdottir A, Jonasdottir A, Oddsson A, Helgason A, Magnusson OT, Walters GB, Frigge ML, Helgadottir HT, Johannsdottir H, Bergsteinsdottir K, Ogmundsdottir MH, Center JR, Nguyen TV, Eisman JA, Christiansen C,

Steingrimsson E, Jonasson JG, Tryggvadottir L, Eyjolfsson GI, Theodors A, Jonsson T, Ingvarsson T, Olafsson I, Rafnar T, Kong A, Sigurdsson G, Masson G, Thorsteinsdottir U, Stefansson K. 2013. Nonsense mutation in the LGR4 gene is associated with several human diseases and other traits. *Nature*, **497**(7450): 517−520.

Thackray LB, Handley SA, Gorman MJ, Poddar S, Bagadia P, Briseño CG, Theisen DJ, Tan Q, Hykes BL, Jr., Lin H, Lucas TM, Desai C, Gordon JI, Murphy KM, Virgin HW, Diamond MS. 2018. Oral antibiotic treatment of mice exacerbates the disease severity of multiple flavivirus infections. *Cell Reports*, **22**(13): 3440−3453.e6.

Van den Bergh B, Michiels JE, Wenseleers T, Windels EM, Boer PV, Kestemont D, De Meester L, Verstrepen KJ, Verstraeten N, Fauvart M, Michiels J. 2016. Frequency of antibiotic application drives rapid evolutionary adaptation of escherichia coli persistence. *Nature Microbiology*, **1**(9): 16020.

Vandamme TF. 2015. Rodent models for human diseases. *European Journal of Pharmacology*, **759**: 84−89.

Wienhold SM, Macrì M, Nouailles G, Dietert K, Gurtner C, Gruber AD, Heimesaat MM, Lienau J, Schumacher F, Kleuser B, Opitz B, Suttorp N, Witzenrath M, Müller-Redetzky HC. 2018. Ventilator-induced lung injury is aggravated by antibiotic mediated microbiota depletion in mice. *Critical Care*, **22**(1): 282.

Wills-Karp M, Luyimbazi J, Xu X, Schofield B, Neben TY, Karp CL, Donaldson DD. 1998. Interleukin-13: Central mediator of allergic asthma. *Science*, **282**(5397): 2258−2261.

Wu S, Jiang ZY, Sun YF, Yu B, Chen J, Dai CQ, Wu XL, Tang XL, Chen XY. 2013. Microbiota regulates the TLR7 signaling pathway against respiratory tract influenza a virus infection. *Current Microbiology*, **67**(4): 414−422.

Yang LY, Nossa CW, Pei ZH. 2015. Microbial dysbiosis and esophageal diseases. *In* : Highlander SK, Rodriguez-Valera F, White BA (eds). Encyclopedia of Metagenomics. Boston, MA: Springer, 379-384, https://doi.org/10.1007/978-1-4899-7475-4_65.

Yeruva L, Spencer NE, Saraf MK, Hennings L, Bowlin AK, Cleves MA, Mercer K, Chintapalli SV, Shankar K, Rank RG, Badger TM, Ronis MJ. 2016. Formula diet alters small intestine morphology, microbial abundance and reduces VE-cadherin and IL-10 expression in neonatal porcine model. *BMC Gastroenterology*, **16**: 40.

Yousefi B, Eslami M, Ghasemian A, Kokhaei P, Farrokhi AS, Darabi N. 2019. Probiotics importance and their immunomodulatory properties. *Journal of Cellular Physiology*, **234**(6): 8008−8018.

Increased attention to snake images in cynomolgus monkeys: an eye-tracking study

Bo Zhang[1,2,*], Zhi-Gang Zhou[2], Yin Zhou[2], Yong-Chang Chen[2,*]

[1] *Guangxi Key Laboratory of Brain and Cognitive Neuroscience, Guilin Medical University, Guilin, Guangxi 541199, China*

[2] *Yunnan Key Laboratory of Primate Biomedicine Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, Yunnan 650500, China*

ABSTRACT

Previous studies have revealed faster detection of snake images in humans and non-human primates (NHPs), suggesting automatic detection of evolutionary fear-relevant stimuli. Furthermore, human studies have indicated that general fear-relevance rather than evolutionary relevance is more effective at capturing attention. However, the issue remains unclarified in NHPs. Thus, in the present study, we explored the attentional features of laboratory-reared monkeys to evolutionary and general fear-relevant stimuli (e.g., images of snakes, capturing gloves). Eye-tracking technology was utilized to assess attentional features as it can provide more accurate latency and variables of viewing duration and frequency compared with visual search task (VST) and response latency adopted in previous studies. In addition, those with autism spectrum disorder (ASD) show abnormal attention to threatening stimuli, including snake images. Rett syndrome (RTT) is considered a subcategory of ASD due to the display of autistic features. However, the attentional features of RTT patients or animal models to such stimuli remain unclear. Therefore, we also investigated the issue in *MECP2* gene-edited RTT monkeys. The influence of different cognitive loads on attention was further explored by presenting one, two, or four images to increase stimulus complexity. The eye-tracking results revealed no significant differences between RTT and control monkeys, who all presented increased viewing (duration and frequency) of snake images but not of aversive stimuli compared with control images, thus suggesting attentional preference for evolutionary rather than general fear-relevant visual stimuli. Moreover, the preference was only revealed in visual tasks composed of two or four images, suggesting its cognitive-load dependency.

Keywords: Non-human primates; Attention; Snake; Evolutionary relevance

INTRODUCTION

Animals can acquire fear response to neutral stimuli through associative learning. For instance, laboratory-reared rodents can learn fear response to neutral sounds/environments through fear conditioning (Maren, 2001). However, animals demonstrate instant fear to evolutionary-relevant predators. For example, rodents display fear response to cat odor despite never having encountered a live cat, suggesting the

development of innate fear to natural predators over eons of evolution (Dielenberg & McGregor, 2001). In contrast, early non-human primate (NHP) studies by Mineka et al. (1980) have suggested the requirement of associative learning for the development of fear response to snakes, a primary predator of NHPs from an evolutionary perspective (Isbell, 2006; Öhman, 2009). For instance, their studies have indicated that wild monkeys demonstrate intense snake fear, probably due to encountering snakes in the wild, whereas laboratory-reared monkeys do not (Cook & Mineka, 1990; Mineka et al., 1980). However, laboratory-reared monkeys have been shown to acquire a preferential fear response to toy snakes rather than to flowers by observational learning, suggesting differential perception of evolutionary fear-relevant and -irrelevant stimuli (Cook & Mineka, 1989, 1990).

Earlier research proposed an evolved module for fear learning, suggesting automatic detection and unconscious perception of evolutionary fear-relevant stimuli by a primitive evolved subcortical brain network (Öhman & Mineka, 2001, 2003). Subsequent human and NHP studies indeed revealed faster detection of snake images compared with images of flowers or mushrooms using the visual search task (VST), thus supporting the fear learning module (Blanchette, 2006; Kawai & Koda, 2016; Shibasaki & Kawai, 2009; Soares et al., 2014). However, faster detection of general fear-relevant stimuli, such as knifes, guns, and syringes, has also been revealed in humans, suggesting more importance of general fear-relevance than evolutionary relevance in capturing attention (Blanchette, 2006; Brown et al., 2010; Forbes et al, 2011). With the issue remaining unclarified in NHPs, we explored the attentional features of laboratory-reared NHPs to general fear-relevant stimuli (e.g., images of capturing gloves, nets) in addition to evolutionary fear-relevant stimuli (e.g., images of snakes). In regard to the accuracy of response latency, the stimulus detection speed in previous VST studies is inevitably affected by the additional body and hand responses after the initial target detection. Thus, in the current study, we utilized eye-tracking, a widely adopted technology in previous NHP attentional studies, as it provides more accurate latency and additional variables of attention, such as viewing duration and frequency (Dal Monte et al., 2015; Gothard et al., 2004; Machado et al., 2011; Zhang et al., 2012).

Rett syndrome (RTT), which is considered a category of autism spectrum disorder (ASD) due to its display of autistic features, exhibits unique attentional patterns (Rose et al., 2013, 2016). Although ASD patients demonstrate abnormal attention to threatening stimuli, including snake images (Isomura et al., 2015; Milosavljevic et al., 2017), their attentional features to such stimuli remain unclear. Changes in methyl-CpG-binding protein 2 (*MECP2*) gene expression are associated with neurodevelopmental disorders, such as ASD and RTT (Qiu, 2018). Therefore, we established the first *MECP2* gene mutant RTT monkey model to explore the cognitive phenotypes and potential behavioral interventions of the syndrome (Chen et al., 2017). We recently reported increased attention of RTT monkeys to salient social stimuli (conspecific stare and profile faces) compared with control monkeys, indicating social valence-related attentional preference (Zhang et al., 2019b). Thus, we further investigated the attentional features of RTT monkeys to evolutionary-related or general fear-relevant visual stimuli.

MATERIALS AND METHODS

Subjects
Seven laboratory-reared control and five gene-edited RTT cynomolgus monkeys (*Macaca fascicularis*, females, 26–49 months old, weighing 2.2–3.9 kg) were used in the present study. All were born and raised at the Yunnan Key Laboratory of Primate Biomedical Research Center, Kunming, China. They were housed in a controlled environment during the experiment (temperature: 22±1 °C ; humidity: 50%±5% RH) under a 12 h light/12 h dark cycle (lights on at 0800h). All animals were feed a commercial monkey diet twice a day, with fruit and vegetable enrichment provided once a day and with tap water provided *ad libitum*. The animal facility is accredited by AAALAC International and all experimental and animal care procedures were carried out in accordance with the National Institutes of Health guide for the care and use of laboratory animals. All experimental protocols were approved by the Institutional Animal Care and Use Committee of the Yunnan Key Laboratory of Primate Biomedical Research.

Eye-tracking experiment
The noninvasive head restraint method for NHP eye-tracking has been described in our previous reports (Zhang et al., 2012, 2019b). Briefly, the restraint helmet was made with thermoplastic mesh sheets (30 cm wide×30 cm long×0.32 cm thick) bordered on four sides by a hard, plastic frame, through which the helmet was attached to the primate chair with wing bolts. The helmet was made pliable by soaking in warm water, then stretched by the experimenter and gently placed and molded closely over the monkey's head (subjects were sedated with ketamine hydrochloride, 5–10 mg/kg , im). The front part of the helmet was then cut to expose the eye and snout regions once the thermoplastic sheet became rigid after cooling.

The detailed eye-tracking procedures, including the apparatus setup, behavioural adaptation, and pre-test calibration, have been described previously (Zhang et al., 2019b). Briefly, a Tobii Pro TX300 Eye Tracker (Tobii Technology AB, Danderyd, Sweden) was set in front of each monkey at a distance of 65 cm, with infrared illumination and cameras used to induce and detect pupil and corneal reflections, respectively, which were then integrated with an internal model to enable calculation of gaze data. Gaze data were sampled at 300 Hz and integrated with a 23 inch monitor at a resolution of 1 920×1 080 pixels, with a single fixation then registered for each gaze point falling within a 0.5° visual angle radius over 75 ms. Visual stimulus presentation and gaze data were generated and collected on a PC with Tobii Studio software (v3.4.7). Monkeys were gradually habituated to the testing environment and manipulations in the testing

room with lights turned off to diminish distraction. Pre-test calibration was performed by presenting small circles or audio cartoon symbols at pre-set locations to attract attention. Results were then presented graphically on the screen, with a small green dot in the center of a circle representing qualified high-quality calibration (Figure 1C). Low-quality calibration resulted in green colored lines extending from the circle, with the number, length, and dispersion of the lines representing the extent of mismatch between acquired data and actual calibration image location. Acceptance of calibration results was determined by visual inspection by an experienced experimenter. The calibration routine was repeated to attain satisfactory results within 5 min, or else the task was terminated for the subject.
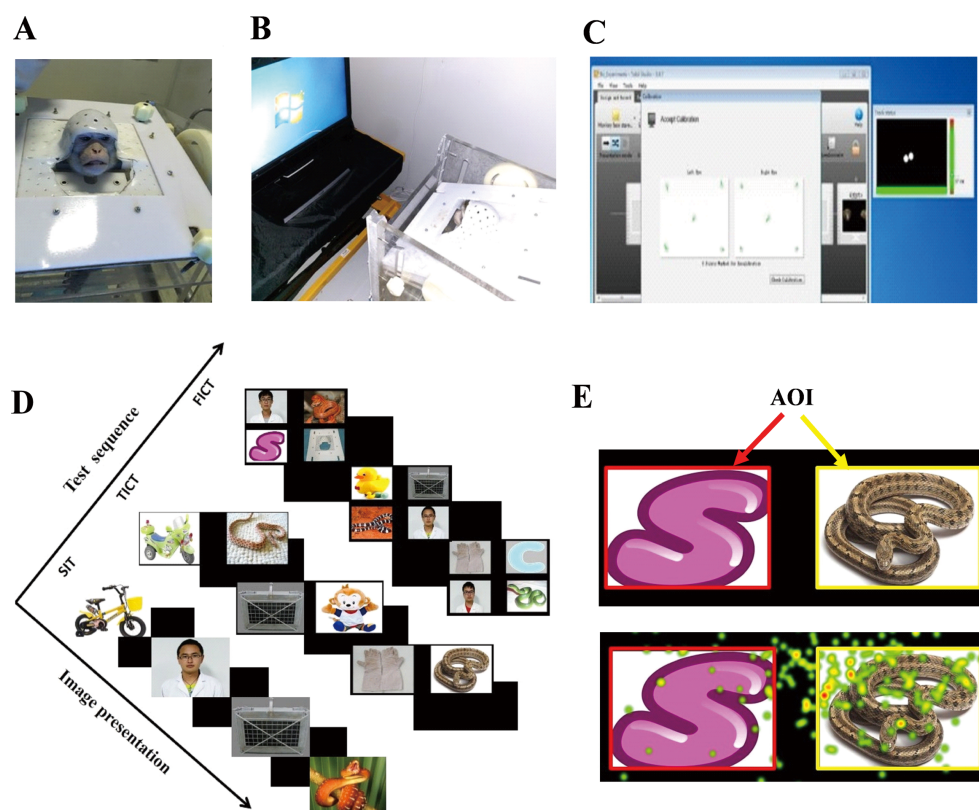

Figure 1 Demonstrations of the eye-tracking apparatus, procedure and visual stimuli

A: Non-invasive head restraint method. B: Behavioral eye-tracking setup, in which a conscious monkey wearing a restraint helmet sits in a primate chair and faces the Tobii Pro TX300 Eye Tracker. C: Five-point calibration results for both eyes indicated by green dots. D: Description of three tasks performed sequentially, including single image test (SIT), two image comparison test (TICT), and four image comparison test (FICT). E: Top panel, example of manually drawn areas of interest (AOIs, red and yellow squares); bottom panel, heatmap example of eye-tracking results within different AOIs, with red representing most attended area. B and C are quoted and modified from our previous report (Zhang et al., 2019b).

Four categories of visual stimuli were used, including evolutionary fear-relevant stimuli (snakes), general fear-relevant aversive stimuli (capturing gloves, nets), control stimuli (toys, such as ducks and hello kitty; abstract graphics, such as letter S in Figure 1E) and familiar human neutral faces. Images of conspecific faces were used in previous investigation and, therefore, were not included in the present study so as to avoid the potential influence of stimulus overexposure (Zhang et al., 2019b). To explore the influence of cognitive load on attention, increased complexity of visual stimuli was introduced by presenting one image, two paired images, and then four paired images on the screen across paradigms of single image test (SIT), two image comparison test (TICT), and four image comparison test (FICT),

respectively. For SIT, eight images were used in each category (dimension 570×380 pixels). For TICT, three categories of paired images were used, including aversive stimuli-control, aversive stimuli-snake, and snake-control, with eight pairs for each category (dimension 1 700×550 pixels). For FICT, 10 paired images were used (dimension 1 700×1 050 pixels), with each image composed of one picture from each category. Each image was presented once consequentially, with 32 trials in SIT, 24 trials in TICT, and 10 trials in FICT. The sequence of tasks was SIT, TICT, and then FICT, with each task performed once (Figure 1D). The order of presented images was randomized by Tobii software and the presentation duration of each image was 5 s in SIT, 8 s in TICT, and 15 s in FICT, with 2 s of blank image as an inter-

trial interval.

Data analyses

Rectangular areas of interest (AOIs) were hand drawn by the analyzer on each image (Figure 1E), with Tobii Studio software (v3.4.7) automatically calculating total viewing duration, frequency, and latency within different AOIs. All variables were subject to square root transformation to generate normally distributed data for further statistical analyses using SPSS 19.0 (SPSS Inc, Chicago, Illinois, USA). The data were assessed by analysis of variance (ANOVA) with repeated measures, with groups being the between-subjects factor and stimuli being the within-subjects factor. All data were presented as means±SEM and differences were considered significant at $P<0.05$.

RESULTS

Two-way repeated ANOVA revealed no significant interaction between group and stimuli in SIT (duration $F_{(3,30)}$=0.157, frequency $F_{(3,30)}$ =0.189, latency $F_{(3,30)}$ =0.513), snake-control comparison (duration $F_{(1,10)}$ =0.288, frequency $F_{(1,10)}$=0.01, latency $F_{(1,10)}$ =3.452), and FICT (duration $F_{(3,30)}$=1.186, frequency $F_{(3,30)}$ =1.35, latency $F_{(3,30)}$ =1.104) (all $P>0.05$). Therefore, the main effects of group and stimuli were further explored. ANOVA also revealed no significant group effect in SIT (duration $F_{(1,10)}$ =0.33, frequency $F_{(1,10)}$ =0.78, latency $F_{(1,10)}$=0), snake-control comparison (duration $F_{(1,10)}$=0.328, frequency $F_{(1,10)}$ =0.21, latency $F_{(1,10)}$ =0.042), and FICT (duration $F_{(1,10)}$ =0.001, frequency $F_{(1,10)}$ =0.01, latency $F_{(1,10)}$=0.069) (all $P>0.05$). Results indicated similar attentional features between RTT and control monkeys, independent of differential cognitive load. However, ANOVA revealed a significant stimulus effect in the snake-control comparison (duration $F_{(1,10)}$ =5.182, frequency $F_{(1,10)}$ =6.388) and FICT (duration $F_{(3,30)}$ =3.099, frequency $F_{(3,30)}$ =3.356) (all $P<0.05$). The subjects presented significantly increased viewing (duration and frequency) of snakes compared with control images in the snake-control comparison (Figure 2D, E) and FICT (*post hoc* comparison, duration P =0.022, frequency P=0.015) (Figure 2G, H). In contrast, ANOVA revealed no significant stimulus effect in SIT (duration $F_{(3,30)}$=1.805, frequency $F_{(3,30)}$ =1.813, all P >0.05) (Figure 2A, B). Additionally, results showed significantly reduced latency viewing of snakes compared with control images in the snake-control image comparison ($F_{(1,10)}$=10.165, P =0.01) (Figure 2F), but not in SIT or FICT (Figure 2C, I), suggesting task-dependent faster detection of snake images. Thus, these results suggest task-dependent attentional preference for snake images compared with control images in both groups of monkeys.

In addition to evolutionary fear-relevant snakes, the present study also included general fear relevant aversive stimuli, such as capturing gloves and nets, which the subjects have encountered and presented fear responses to in daily life. Two-way ANOVA revealed no significant differences between aversive and control image viewing in SIT and FICT (all $P>$

0.05) (Figure 2A, B, G, H) and TICT (duration $F_{(1,10)}$=2.5, frequency $F_{(1,10)}$ =4.735, all $P>0.05$) (Figure 3A, B), indicating no attentional preference for aversive stimuli compared with control images. Together with the snake attentional preference findings, these results suggest attentional preference for evolutionary fear-relevant rather than general fear-relevant stimuli in both groups of monkeys. However, direct comparison of snake-aversive stimuli revealed no significant viewing difference (duration $F_{(1,10)}$ =2.025, frequency $F_{(1,10)}$=1.435, latency $F_{(1,10)}$ =0.003, all $P>0.05$) (Figure 3D–F), indicating task/comparison dependency of the preference.

DISCUSSION

Compared with non-primate animals, NHPs share considerably high similarities with humans in various aspects, such as their genomes and highly developed brains, which make them ideal models for translational medical research (Izpisua Belmonte et al., 2015; Kaas, 2013; Wu et al., 2017; Zhang, 2017, 2019a). For instance, macaque monkeys (genus *Macaca*) are suitable for visual cognition research as they possess trichromatic color vision and high visual acuity (Jacobs, 2008; Orban et al., 2004). Snakes are primary predators of NHPs from an evolutionary perspective (Isbell, 2006; Öhman, 2009), and previous studies have indicated NHPs can acquire fear of snakes via associative learning (Cook & Mineka, 1990; Mineka et al., 1980). Subsequent studies have also found behavioral inhibition (Nelson et al., 2003) and faster detection of snake images (Kawai & Koda, 2016; Shibasaki & Kawai, 2009) in laboratory-reared monkeys, consistent with the evolved fear learning module, which suggests automatic detection of evolutionary fear-relevant snakes (Öhman & Mineka, 2001, 2003).

Related attentional studies have mainly utilized the VST paradigm and consequential behavioral response latency to assess visual attention, with accuracy inevitably affected due to the combination of both hand and body responses (Blanchette, 2006; Kawai & Koda, 2016; Shibasaki & Kawai, 2009; Soares et al., 2014). In contrast, we explored snake attention in laboratory-reared monkeys based on eye-tracking, which provides more accurate attentional latency and additional variables of viewing duration and frequency. In addition to evolutionary fear-relevant snake attention, faster detection of general fear-relevant stimuli has been revealed in humans (Blanchette, 2006; Brown et al., 2010; Forbes et al., 2011; Smith et al., 2003). However, the attentional features to general fear-relevant stimuli remain to be elucidated in NHPs. As such, we included general fear-relevant images, such as capturing gloves and nets, to explore the issue in laboratory-reared monkeys. We recently established the first *MECP2* gene mutant RTT monkey model to explore the visual cognitive phenotypes associated with the syndrome and found that social valence increased attention in the monkeys (Zhang et al., 2019b). In the present study, we further investigated the attentional features of RTT monkeys to evolutionary and general fear-relevant stimuli.

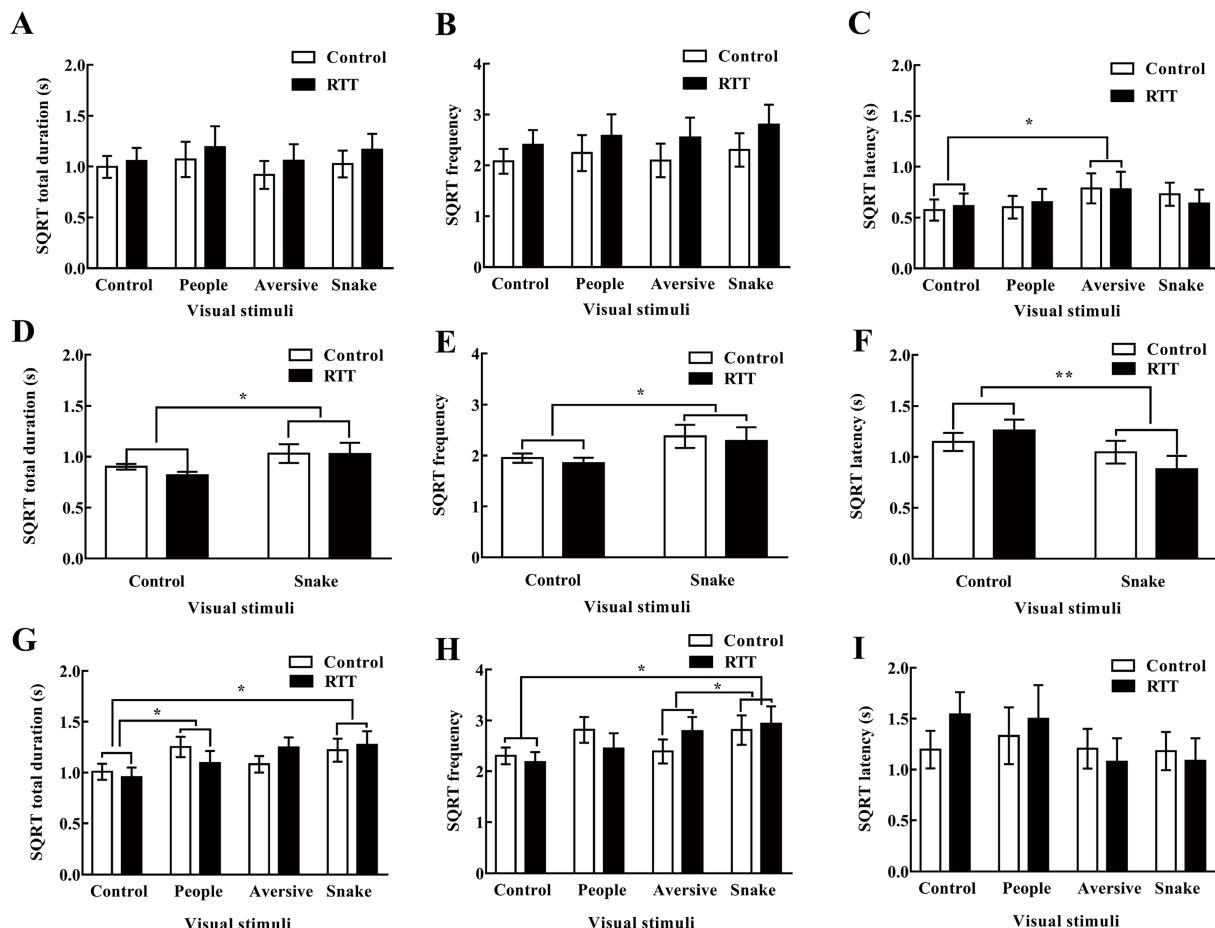We found similar attentional features between RTT and

Figure 2 Single image test with variables of viewing duration (A), frequency (B), and latency (C). Snake-control comparison with variables of viewing duration (D), frequency (E), and latency (F). Four image comparison with variables of viewing duration (G), frequency (H), and latency (I)

All variables were subjected to square root transformation and represented as means±*SEM*. SQRT: Square root transformation, *: $P<0.05$, **: $P<0.01$.

control monkeys, both demonstrating increased viewing duration and frequency of snake images but not of aversive stimuli compared with control images (Figure 2D, E, G, H; Figure 3A, B). The monkeys also demonstrated faster detection of snake images and delayed detection of aversive stimuli, as indicated by decreased (Figure 2F) and increased (Figure 2C, P =0.03) latency, respectively. These results suggest attentional preference (increased viewing and faster detection) for evolutionary related rather than general fear-relevant stimuli, providing the first NHP eye-tracking evidence for the previously proposed evolved fear learning module (Öhman & Mineka, 2001). This null preference for general fear-relevant aversive stimuli is inconsistent with the faster detection findings in humans (Blanchette, 2006; Brown et al., 2010; Forbes et al., 2011; Smith et al., 2003). This inconsistency may be caused by differences in testing paradigms, as VST requires additional hand touching response after the initial visual detection of the target image compared with the eye-tracking procedure. Alternatively, this inconsistency may be due to the unique ability of humans to

use oral/written language for information processing and transmission. For instance, humans can learn that general fear-relevant stimuli (e.g., knife, gun) are threatening from various indirect sources such as parents, books, and news reports. In contrast, animals can only acquire fear response to stimuli through direct experience or observational learning.

We utilized images with different complexity to explore the potential influence of cognitive load on visual attention. Results revealed increased viewing and faster detection of snake images in TICT and FICT (Figure 2), and delayed detection of aversive stimuli in SIT (Figure 2C, latency P=0.03), suggesting cognitive load-affected attention. The SIT only requires passive viewing, whereas TICT and FICT require advanced selective attention and comparison. Thus, this may contribute to the subjects demonstrating increased viewing duration and frequency across SIT (duration 1 s, frequency two counts), TICT (duration 2 s, frequency four counts), and FICT (duration 4 s, frequency eight counts). Alternatively, the increased viewing may be caused by increased duration of image presentation across tests (5 s in SIT, 8 s in TICT, and
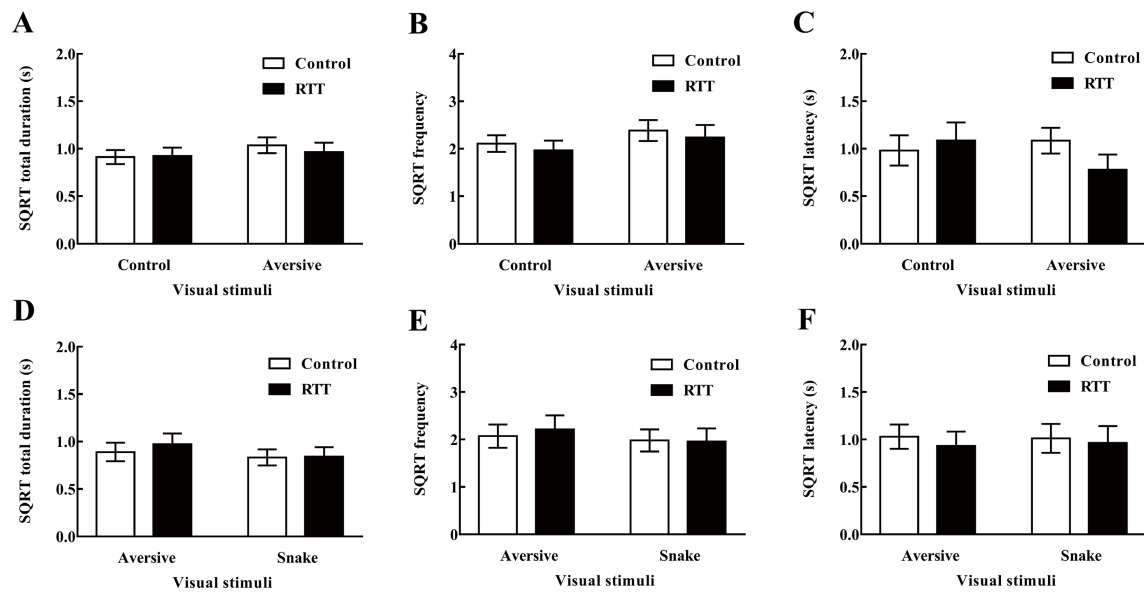
Figure 3 Aversive-control comparison with variables of viewing duration (A), frequency (B), and latency (C). Aversive-snake comparison with variables of viewing duration (D), frequency (E), and latency (F)

All variables were subjected to square root transformation and represented as means±*SEM*. SQRT: Square root transformation.

15 s in FICT). We also found increased viewing (duration 1 s, frequency two counts) of single images presented in SIT (Figure 2A, B) compared with our previous findings (duration 0.6 s, frequency one count) (Zhang et al., 2012). This difference could potentially be caused by different variables, e.g., subjects (female cynomolgus monkeys vs. male rhesus monkeys), visual stimuli (four categories vs. conspecific profile faces), testing apparatus (Tobii Pro TX300 vs. T60 Eye Tracker), and definition of one fixation (gaze maintained within 0.5° visual angle for 75 vs. 100 ms).

Consistent with the present eye-tracking findings of snake attentional preference, previous studies have provided electrophysiological evidence of faster detection of snake images in the pulvinar neurons of macaque monkeys (Le et al., 2016; Van Le et al., 2013). However, automatic attention of snakes does not necessarily mean automatic negative emotion (such as fear) (Purkis & Lipp, 2007). Therefore, further studies are required to validate the direct association between snake attentional preference and fear emotion/response. It would be informative to know whether monkeys present fear-associated biological changes during eye-tracking investigations, such as increased blood pressure, heart rhythm, pupil size, and amygdala activities.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

B. Z. designed the experiment, analyzed the data, and prepared the manuscript. Z.G.Z. and Y.Z. performed the experiments. Y.C.C. participated in experimental design and manuscript preparation. All authors read and approved the final version of the manuscript.

REFERENCES

Blanchette I. 2006. Snakes, spiders, guns, and syringes: how specific are evolutionary constraints on the detection of threatening stimuli?. *Quarterly Journal of Experimental Psychology*, **59**(8): 1484−1504.

Brown C, El-Deredy W, Blanchette I. 2010. Attentional modulation of visual-evoked potentials by threat: investigating the effect of evolutionary relevance. *Brain and Cognition*, **74**(3): 281−287.

Chen Y, Yu J, Niu Y, Qin D, Liu H, Li G, Hu Y, Wang J, Lu Y, Kang Y, Jiang Y, Wu K, Li S, Wei J, He J, Wang J, Liu X, Luo Y, Si C, Bai R, Zhang K, Liu J, Huang S, Chen Z, Wang S, Chen X, Bao X, Zhang Q, Li F, Geng R, Liang A, Shen D, Jiang T, Hu X, Ma Y, Ji W, Sun YE. 2017. Modeling rett syndrome using talen-edited MECP2 mutant cynomolgus monkeys. *Cell*, **169**(5): 945−955.e10.

Cook M, Mineka S. 1989. Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of Abnormal Psychology*, **98**(4): 448−459.

Cook M, Mineka S. 1990. Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of Experimental Psychology Animal Behavior Processes*, **16**(4): 372−389.

Dal Monte O, Costa VD, Noble PL, Murray EA, Averbeck BB. 2015. Amygdala lesions in rhesus macaques decrease attention to threat. *Nature Communications*, **6**: 10161.

Dielenberg RA, Mcgregor IS. 2001. Defensive behavior in rats towards predatory odors: a review. *Neuroscience and Biobehavioral Reviews*, **25**(7−8): 597−609.

Forbes SJ, Purkis HM, Lipp OV. 2011. Better safe than sorry: simplistic fearrelevant stimuli capture attention. *Cognition and Emotion*, **25**(5): 794−804.

Gothard KM, Erickson CA, Amaral DG. 2004. How do rhesus monkeys (Macaca mulatta) scan faces in a visual paired comparison task?. *Animal Cognition*, **7**(1): 25−36.

Isbell LA. 2006. Snakes as agents of evolutionary change in primate brains. *Journal of Human Evolution*, **51**(1): 1−35.

Isomura T, Ogawa S, Shibasaki M, Masataka N. 2015. Delayed disengagement of attention from snakes in children with autism. *Frontiers in Psychology*, **6**: 241.

Izpisua Belmonte JC, Callaway EM, Caddick SJ, Churchland P, Feng G, Homanics GE, Lee KF, Leopold DA, Miller CT, Mitchell JF, Mitalipov S, Moutri AR, Movshon JA, Okano H, Reynolds JH, Ringach D, Sejnowski TJ, Silva AC, Strick PL, Wu J, Zhang F. 2015. Brains, genes, and primates. *Neuron*, **86**(3): 617−631.

Jacobs GH. 2008. Primate color vision: a comparative perspective. *Visual Neuroscience*, **25**(5–6): 619−633.

Kaas JH. 2013. The evolution of brains from early mammals to humans. *Wiley Interdisciplinary Reviews Cognitive Science*, **4**(1): 33−45.

Kawai N, Koda H. 2016. Japanese monkeys (*Macaca fuscata*) quickly detect snakes but not spiders: evolutionary origins of fear-relevant animals. *Journal of Comparative Psychology*, **130**(3): 299−303.

Le QV, Isbell LA, Matsumoto J, Le VQ, Nishimaru H, Hori E, Maior RS, Tomaz C, Ono T, Nishijo H. 2016. Snakes elicit earlier, and monkey faces, later, gamma oscillations in macaque pulvinar neurons. *Scientific Reports*, **6**: 20595.

Machado CJ, Bliss-Moreau E, Platt ML, Amaral DG. 2011. Social and nonsocial content differentially modulates visual attention and autonomic arousal in Rhesus macaques. *PLoS One*, **6**(10): e26598.

Maren S. 2001. Neurobiology of Pavlovian fear conditioning. *Annual Review of Neuroscience*, **24**: 897−931.

Milosavljevic B, Shephard E, Happe FG, Johnson MH, Charman T. 2017. Anxiety and attentional bias to threat in children at increased familial risk for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, **47**(12): 3714−3727.

Mineka S, Keir R, Price V. 1980. Fear of snakes in wild- and laboratory-reared rhesus monkeys (*Macaca mulatta*). *Animal Learning & Behavior*, **8**(4): 653−663.

Nelson EE, Shelton SE, Kalin NH. 2003. Individual differences in the responses of naive rhesus monkeys to snakes. *Emotion*, **3**(1): 3−11.

Öhman A. 2009. Of snakes and faces: an evolutionary perspective on the psychology of fear. *Scandinavian Journal of Psychology*, **50**(6): 543−552.

Öhman A, Mineka S. 2001. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, **108**(3): 483−522.

Öhman A, Mineka S. 2003. The malicious serpent: snakes as a prototypical stimulus for an evolved module of fear. *Current Directions in Psychological Science*, **12**(1): 5−9.

Orban GA, Van Essen D, Vanduffel W. 2004. Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, **8**(7): 315−324.

Purkis HM, Lipp OV. 2007. Automatic attention does not equal automatic fear: preferential attention without implicit valence. *Emotion*, **7**(2): 314−323.

Qiu Z. 2018. Deciphering MECP2-associated disorders: disrupted circuits and the hope for repair. *Current Opinion in Neurobiology*, **48**: 30−36.

Rose SA, Djukic A, Jankowski JJ, Feldman JF, Fishman I, Valicenti-Mcdermott M. 2013. Rett syndrome: an eye-tracking study of attention and recognition memory. *Developmental Medicine & Child Neurology*, **55**(4): 364−371.

Rose SA, Djukic A, Jankowski JJ, Feldman JF, Rimler M. 2016. Aspects of attention in Rett Syndrome. *Pediatric Neurology*, **57**: 22−28.

Shibasaki M, Kawai N. 2009. Rapid detection of snakes by Japanese monkeys (Macaca fuscata): an evolutionarily predisposed visual system. *Journal of Comparative Psychology*, **123**(2): 131−135.

Smith NK, Cacioppo JT, Larsen JT, Chartrand TL. 2003. May I have your attention, please: electrocortical responses to positive and negative stimuli. *Neuropsychologia*, **41**(2): 171−183.

Soares SC, Lindstrom B, Esteves F, Öhman A. 2014. The hidden snake in the grass: superior detection of snakes in challenging attentional conditions. *PLoS One*, **9**(12): e114724.

Van Le Q, Isbell LA, Matsumoto J, Nguyen M, Hori E, Maior RS, Tomaz C, Tran AH, Ono T, Nishijo H. 2013. Pulvinar neurons reveal neurobiological evidence of past selection for rapid detection of snakes. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(47): 19000−19005.

Wu SH, Liao ZX, Rizak JD, Zheng N, Zhang LH, Tang H, He XB, Wu Y, He XP, Yang MF, Li ZH, Qin DD, Hu XT. 2017. Comparative study of the transfection efficiency of commonly used viral vectors in rhesus monkey (Macaca mulatta) brains. *Zoological Research*, **38**(2): 88−95.

Zhang B. 2017. Consequences of early adverse rearing experience(EARE) on development: insights from non-human primate studies. *Zoological Research*, **38**(1): 7−35.

Zhang B, Noble PL, Winslow JT, Pine DS, Nelson EE. 2012. Amygdala volume predicts patterns of eye fixation in rhesus monkeys. *Behavioural Brain Research*, **229**(2): 433−437.

Zhang B, Xiong F, Ma Y, Li B, Mao Y, Zhou Z, Yu H, Li J, Li C, Fu J, Wang J. 2019a. Chronic phencyclidine treatment impairs spatial working memory in rhesus monkeys. *Psychopharmacology*, **236**(7): 2223−2232.

Zhang B, Zhou Z, Zhou Y, Zhang T, Ma Y, Niu Y, Ji W, Chen Y. 2019b. Social valence related increased attention in Rett syndrome (RTT) cynomolgus monkeys: An eye-tracking study. *Autism Research*, **12**(11): 1585−1597.

D-dopachrome tautomerase from Japanese sea bass (*Lateolabrax japonicus*) is a chemokine-like cytokine and functional homolog of macrophage migration inhibitory factor

Feng Xu[1,2], Ming-Yun Li[2], Jiong Chen[1,2,3,*]

[1] *State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Ningbo University, Ningbo, Zhejiang 315211, China*

[2] *Laboratory of Biochemistry and Molecular Biology, School of Marine Sciences, Ningbo University, Ningbo, Zhejiang 315832, China*

[3] *Key Laboratory of Applied Marine Biotechnology of Ministry of Education, Ningbo University, Ningbo, Zhejiang 315832, China*

ABSTRACT

D-dopachrome tautomerase (DDT), a member of the macrophage migration inhibitory factor (MIF) protein superfamily, is a newly described cytokine with chemokine-like characteristics. However, research on fish DDT remains limited. In this study, we identified a DDT homolog (LjDDT) from the Japanese sea bass, *Lateolabrax japonicus*. Sequence analysis showed that LjDDT had typical sequence features of known DDT and MIF homologs and was most closely related to DDT of rock bream (*Oplegnathus fasciatus*). *LjDDT* transcripts were detected in all tested tissues of healthy Japanese sea bass, with the highest expression found in the liver. Upon infection with *Vibrio harveyi*, *LjDDT* transcripts were significantly down-regulated in the three tested tissues, including the liver, spleen, and head kidney. Recombinant LjDDT (rLjDDT) and the corresponding antibody (anti-rLjDDT) were subsequently prepared. The administration of 100 µg/g anti-rLjDDT had a statistically significant protective effect on the survival of *V. harveyi*-infected fish. Moreover, rLjDDT was able to induce the migration of monocytes/macrophages (MO/MΦ) and lymphocytes both *in vitro* and *in vivo*, but without significant influence on the migration of neutrophils. rLjDDT exhibited chemotactic activity for lipopolysaccharide (LPS)-stimulated M1-type MO/MΦ *in vitro*, but not for cAMP-stimulated M2-type MO/MΦ. Furthermore, the knockdown of *LjCD74*, but not *LjCXCR4*, significantly down-regulated the rLjDDT-enhanced migration of MO/MΦ and relieved the rLjMIF-inhibited migration of MO/MΦ. These results indicate that LjCD74 may be the major chemotactic receptor of LjDDT and LjMIF in Japanese sea bass MO/MΦ. Combined rLjDDT+ rLjMIF treatment had no significant effect on the migration of MsiRNA, LjCD74si-, or LjCXCR4-sitreated MO/MΦ compared to the control group, suggesting that the roles of LjDDT and LjMIF may be

antagonistic. In conclusion, our study demonstrates for the first time that DDT may play a role in the immune responses of fish against bacterial infection through chemotactic recruitment of MO/MΦ via mediation of CD74 as an antagonist of MIF.

Keywords: Cell migration; D-dopachrome tautomerase; Japanese sea bass; Macrophage migration inhibitory factor; Monocyte/macrophage

INTRODUCTION

Macrophage migration inhibitory factor (MIF) was first reported to inhibit the random migration of peritoneal lymphocytes and macrophages in hypersensitized guinea pigs (Bloom & Bennett, 1966; David, 1966). It is a pleiotropic proinflammatory cytokine with multiple biological functions in both innate and acquired immunity (Günther et al., 2019). MIF has chemokine-like characteristics (Bernhagen et al., 2007; Sinitski et al., 2019) and also plays a role in pathological diseases, including autoimmune diseases (Rijvers et al., 2018). MIF exerts its biological functions through autocrine and paracrine signaling via binding to and activating its receptors, including HLA class II histocompatibility antigen gamma chain (CD74), C-X-C motif chemokine receptor 4 (CXCR4), and C-X-C motif chemokine receptor 2 (CXCR2) (Bernhagen et al., 2007; Jankauskas et al., 2019; Klasen et al., 2014; Leng & Bucala, 2006; Rajasekaran et al., 2016; Rijvers et al., 2018). For example, MIF promotes the migration of B-cells through a zeta chain of the T-cell receptor-associated protein kinase 70 (ZAP70) - dependent pathway, which is mediated by the cooperative engagement of CXCR4 and CD74 (Klasen et al., 2014).

D-dopachrome tautomerase (DDT), which is a newly described cytokine and a member of the MIF protein superfamily, has attracted increasing research attention (Furukawa et al., 2016; Ma et al., 2019; Merk et al., 2012). DDT was originally identified as an enzyme in the cytoplasm of human melanoma, human liver, and rat organs, which converts D-dopachrome into 5,6-dihydroxyindole (Odh et al., 1993). The DDT gene is related to MIF in terms of sequence, enzyme activity, and gene structure (Esumi et al., 1998; Sugimoto et al., 1999). Human DDT shares 34% amino acid identity with MIF and is located within 80 kb of MIF in genomes (Merk et al., 2012). Recent studies have revealed that DDT is a functional homolog of MIF (Coleman et al., 2008; Merk et al., 2012). In mammals, DDT is associated with numerous physiological processes, including cell recruitment and migration (Rajasekaran et al., 2016; Wang et al., 2017), tumorigenesis and cancer progress (Coleman et al., 2008; Guo et al., 2016; Wang et al., 2017), and inflammatory and autoimmune diseases (Benedek et al., 2017; Fagone et al., 2018; Günther et al., 2019; Kim et al., 2017). DDT also binds to and signals via CD74 but differs from MIF by lacking the pseudo-(E)LR motif necessary for activation of chemokine receptors (Jankauskas et al., 2019; Tilstam et al., 2017; Weber et al., 2008). DDT sequences have been found in many species of fish, but studies on their biological functions are rare (Oh et al., 2013). Recombinant DDT in the rock bream (*Oplegnathus fasciatus*) induces the expression of proinflammatory cytokines such as tumor necrosis factor alpha (TNF-α), interleukin-8 (IL-8), and interleukin-1β (IL-1β) in head kidney cells, indicating that DDT may be involved in the inflammatory responses of fish (Oh et al., 2013).

The Japanese sea bass (*Lateolabrax japonicus*) is a euryhaline marine fish species commonly farmed in China, Japan, and Korea due to its high commercial value. With the growth of the marine aquaculture industry, outbreaks of infectious diseases have become increasingly frequent, leading to serious output declines and economic losses (Zhou et al., 2014). *Vibrio harveyi* has been identified as a major pathogen and cause of vibriosis disease in Japanese sea bass (Zhou et al., 2014). Studying the immune system of marine fish will provide a better understanding of their immune responses to antigenic substances and related mechanisms and may help to develop better disease management strategies for fish farmed under harsh environments. We previously found that Japanese sea bass MIF (LjMIF) can inhibit trafficking of monocytes/macrophages (MO/MΦ) and lymphocytes, enhance phagocytosis and intracellular killing of *V. harveyi* by MO/MΦ, and aggravate *V. harveyi* infection (Xu et al., 2019). In the present study, we identified a Japanese sea bass *DDT* (*LjDDT*) and investigated the relationship between *LjDDT* mRNA expression and *V. harveyi* infection. Moreover, we determined the effects of LjDDT on the regulation of immune cell trafficking and MO/MΦ function *in vitro*. The functional relationships between LjDDT and LjMIF and their receptors LjCD74 and LjCXCR4 were also investigated.

MATERIALS AND METHODS

Fish rearing

Healthy Japanese sea bass, weighing approximately 100 g, were obtained from a commercial farm in Xiangshan County, Ningbo City, China. Fish were maintained in experimental tanks filled with artificial seawater (salinity 20±2, pH 7.5±0.4, temperature 27±1 °C) and acclimated to laboratory conditions for two weeks prior to experimentation. All fish were healthy before the experiment. All experiments were performed in accordance with the Experimental Animal Management Law of China and approved by the Animal Ethics Committee of Ningbo University.

Sequence analysis of *LjDDT*

cDNA sequences of *LjDDT* were retrieved from three newly determined transcriptomes of Japanese sea bass annotated by the Beijing Genomics Institute, China (data not shown). The DDT homolog sequence was then amplified via polymerase chain reaction (PCR) using the cDNA template of Japanese sea bass and authenticated by further cloning, sequencing, and BLAST searching (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The signal peptide was predicted using SignalP v4.1 (http://www.cbs.dtu.dk/services/SignalP/). The

protein domain architecture was analyzed using SMART (http://smart.emblheidelberg.de/). Multiple alignments were carried out using ClustalW (http://clustalw.ddbj.nig.ac.jp/). Non-classical secretion was analyzed using SecretomeP 2.0 (http://www.cbs.dtu.dk/services/SecretomeP/). Phylogenetic and molecular evolutionary analyses were conducted using MEGA v7 (Kumar et al., 2016). The cDNA sequences of *DDT*s or *MIF*s used in this study are listed in Supplementary Table S1.

Tissue mRNA expression analysis of *LjDDT* in Japanese sea bass under healthy and pathological conditions
In vivo bacterial challenge was performed as described previously (Xu et al., 2019). Briefly, the *V. harveyi* strain ATCC33866, which was purchased from the China General Microbiological Culture Collection Center (China), was cultured in Tryptic Soy Broth (TSB) medium at 28 °C with constant shaking at 200 r/min until the logarithmic growth phase. The harvested *V. harveyi* cells were washed three times and resuspended in 100 µL of sterile phosphate buffered saline (PBS). The experimental groups were infected by an intraperitoneal (ip) injection of *V. harveyi* (5×10^6 colony-forming units (CFU) per fish), according to the determined 50% lethal dose (LD_{50}) in 72 h; the same volume of PBS was used for the control group. The liver, spleen, and head kidney were collected from fish at 6, 12, 24, 36, and 48 h post infection (hpi) for pathology-related mRNA expression analysis using quantitative real-time PCR (qRT-PCR). The liver, spleen, head kidney, trunk kidney, gill, intestine, brain, skin, muscle, and heart of healthy Japanese sea bass were also collected for tissue mRNA expression pattern analysis using qRT-PCR.

DNase I digestion and first-strand cDNA synthesis were conducted as reported previously (Chen et al., 2019). Based on the cDNA sequence of *LjDDT* , primers LjDDT-F(+): 5'-AAACCAGAGGACAGGATGAATC-3' and LjDDT-R(-): 5'-CACACCGATAGCAGACACC-3' were designed for the detection of the *LjDDT* transcript by qRT-PCR. Amplification was performed using TB Green Premix Ex Taq II (Takara Bio, Japan), and the reaction mixture was incubated in an ABI StepOne Real-Time PCR System (Applied Biosystems, USA) as follows: 94 °C for 180 s, 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s, followed by melting curve analysis at 94 °C for 30 s, 72 °C for 30 s, and 94 °C for 30 s. Relative expression of *LjDDT* was normalized to that of *Lj18S* rRNA. Samples obtained under healthy and pathological conditions were assessed using the $2^{-\Delta CT}$ and $2^{-\Delta\Delta CT}$ methods, respectively. Each experiment was performed in triplicate and repeated four times.

Prokaryotic expression of LjDDT and antibody preparation
Primer pair LjDDT-p(+): 5'-G<u>GAATTC</u>ATGCCTTTCATCAACT TAGAGAG-3' (underlined section is restriction site for *Eco*R I) and LjDDT-p(-): 5'-G<u>CTCGAG</u>TCACAAGAAGCTCATGACGG T-3' (underlined section is restriction site for *Xho* I) was designed for amplification of the complete open reading frame

(ORF) sequence of LjDDT. After restriction enzyme digestion, the amplicon was cloned into the *Eco*R I/*Xho* I-digested pET-28a (+) expression vector for the construction of plasmid pET-28a-LjDDT. pET-28a-LjDDT was subsequently transformed into the *Escherichia coli* strain BL21 (DE3). The overexpression of recombinant LjDDT (rLjDDT) was induced by isopropyl-β-D-thiogalactopyranoside (IPTG). rLjDDT was purified using a nickel-nitrilotriacetic acid (Ni-NTA) column (QIAGEN, China) at 4 °C. Lipopolysaccharide (LPS) was removed using Detoxi-Gel (Thermo Fisher Scientific, USA). The purified rLjDDT was then used as an antigen to immunize mice to produce antiserum. The anti-rLjDDT IgG (anti-rLjDDT) and isotype IgG (IsoIgG) were purified from mouse sera using Protein G HP SpinTrap columns (GE Healthcare, USA) and their concentrations were determined using the Bradford protein assay. The specificity of the antibody was tested by Western blotting and visualized using an enhanced chemiluminescence (ECL) kit (Advansta, USA), as described previously (Ren et al., 2019). The lyophilized rLjDDT and anti-rLjDDT were kept at –20 ºC until use.

Fish survival assay
Healthy fish were randomly divided into eight groups for survival study: i.e., (1) Control, ip-injected with 100 µL of PBS 30 min post *V. harveyi* (1×10^4 CFU/fish) infection; (2) ip-injected with rLjDDT (1 µg/g body weight) 30 min post *V. harveyi* (1×10^4 CFU/fish) infection; (3) ip-injected with rLjDDT (10 µg/g body weight) 30 min post *V. harveyi* (1×10^4 CFU/fish) infection; (4) ip-injected with rLjDDT (100 µg/g body weight) 30 min post *V. harveyi* (1×10^4 CFU/fish) infection; (5) ip-injected with anti-rLjDDT (1 µg/g body weight) 1 h before *V. harveyi* (1×10^4 CFU/fish) infection; (6) ip-injected with anti-rLjDDT (10 µg/g body weight) 1 h before *V. harveyi* (1×10^4 CFU/fish) infection; (7) ip-injected with anti-rLjDDT (100 µg/g body weight) 1 h before *V. harveyi* (1×10^4 CFU/fish) infection; and (8) ip-injected with IsoIgG (10 µg/g body weight) 1 h before *V. harveyi* (1×10^4 CFU/fish) infection. Over the next 9 d, the fish were monitored daily for death or moribund state. The Kaplan-Meier method was used to analyze the 9 d survival rate.

Isolation of MO/MΦ, lymphocytes, and neutrophils from peripheral blood
MO/MΦ, lymphocytes, and neutrophils were separated from caudal vein blood of healthy Japanese sea bass according to a previously described method (Liu et al., 2018). Briefly, heparinized blood was collected, and cells were isolated following sedimentation with 6% dextran T 500 (Sigma, USA). After low-speed centrifugation at 400 *g* for 25 min at 24 °C, cells packed below Ficoll-Hypaque PREMIUM (GE Healthcare) (i. e., erythrocytes and neutrophils) were subjected to hypotonic lysis with ice-cold ACK (Ammonium-Chloride-Potassium) Lysis Buffer (0.15 mol/L NH_4Cl, 0.01 mol/L $KHCO_3$, 0.1 m mol/L EDTA) to eliminate red blood cells. The resulting neutrophil suspension was washed and suspended in RPMI 1640 medium (Invitrogen, China). The buffer layer above the Ficoll-Hypaque PREMIUM was collected and washed carefully, and the number of cells was

determined using a hemocytometer (Sangon, China). Cells were cultured in 35 mm dishes for 12 h, and adherent MO/MΦ and non-adherent lymphocytes were carefully collected and cultured in complete medium (RPMI 1640 supplemented with 5% (v/v) Japanese sea bass serum, 5% (v/v) fetal bovine serum (FBS, Invitrogen), 100 U/mL penicillin, and 100 μg/mL streptomycin) at 24 °C with 5% CO_2.

Primary culture of Japanese sea bass head kidney-derived MO/MΦ

Head kidney-derived MO/MΦ were isolated from healthy Japanese sea bass and cultured as described previously (Chen et al., 2014). Briefly, leukocyte-enriched fractions were obtained from the Ficoll-medium interface using a Ficoll density gradient (1.077±0.001 g/mL) (Invitrogen) and seeded into 35 mm dishes. After overnight incubation at 24 °C, non-adherent cells were removed by washing and adherent cells were subsequently cultured in complete medium at 24 °C with 5% CO_2.

In vitro cell migration assay

In vitro cell migration assay was performed in a 24-well Transwell chamber (Corning, USA). For the assay of peripheral blood-derived cells, rLjDDT or rLjMIF in complete medium was added to the lower chambers at concentrations of 0, 1.0, and 10.0 μg/mL respectively; MO/MΦ, neutrophils, or lymphocytes were plated on the upper chambers. The chambers were incubated at 24 °C for 4 h. Cells that migrated from the upper to lower chambers were counted using light microscopy (Nikon, Japan). Each migration assay was performed in quadruplicate.

For the polarized MO/MΦ assay, the isolated Japanese sea bass head kidney-derived MO/MΦ were treated with 10.0 μg/mL LPS or 0.5 mg/mL cyclic adenosine monophosphate (cAMP) for 12 h to produce M1 or M2 type MO/MΦ, as described previously (Chen et al., 2018). The *in vitro* chamber assay was then used to determine the chemotactic effect of rLjDDT (at concentrations of 0, 1.0, and 10.0 μg/mL, respectively) on M1 and M2 MO/MΦ, with non-stimulated MO/MΦ used as the control.

In vivo cell migration assay

Fish in the experimental groups were ip-injected with 1.0 μg/g or 10.0 μg/g rLjDDT or rLjMIF per fish in 100 μL of PBS; fish in the control group received the same volume of PBS. Peritoneal cells were collected at 24 hpi and rinsed with 2 mL of PBS using a single-use aseptic injector. After centrifugation at 800 *g* for 8 min at 24 °C, cell pellets were retained and resuspended in 1 mL of PBS. The direct cell counts were evaluated at 400× magnification using a hemocytometer. MO/MΦ, lymphocytes, and neutrophils were further identified microscopically via the Wright-Giemsa staining technique according to previously described methods (Yu et al., 2019).

In vitro MO/MΦ migration after *LjCD74* and *LjCXCR4* knockdown

Japanese sea bass head kidney-derived MO/MΦ were transfected with LjCD74 (MK605507) small interfering RNA

(LjCD74si) (5'-GCUCCAAUGAGGAUGCAAATT-3') or LjCXCR4 (MK605474) siRNA (LjCXCR4si) (5'-CCAACACUCCAGGAUCAUUTT-3') for 48 h to knock down the expression of the target gene, with Mismatched siRNA (MsiRNA) (5'-UUCUCCGAACGUGUCACGUTT-3') treatment used as the negative control. qRT-PCR was used to confirm knockdown of LjCD74 and LjCXCR4 expression. These MO/MΦ were then plated on the upper chambers, and rLjDDT (or rLjMIF) in complete medium was added to the lower chambers at a concentration of 10.0 μg/mL. The *in vitro* cell migration assay was performed as described in the previous section.

Statistical analysis

All data are presented as mean±standard error of mean (*SEM*). Statistical analysis was performed using one-way analysis of variance (ANOVA) with SPSS v13.0 (SPSS Inc., Chicago, USA). A *P*- value of <0.05 were considered statistically significant.
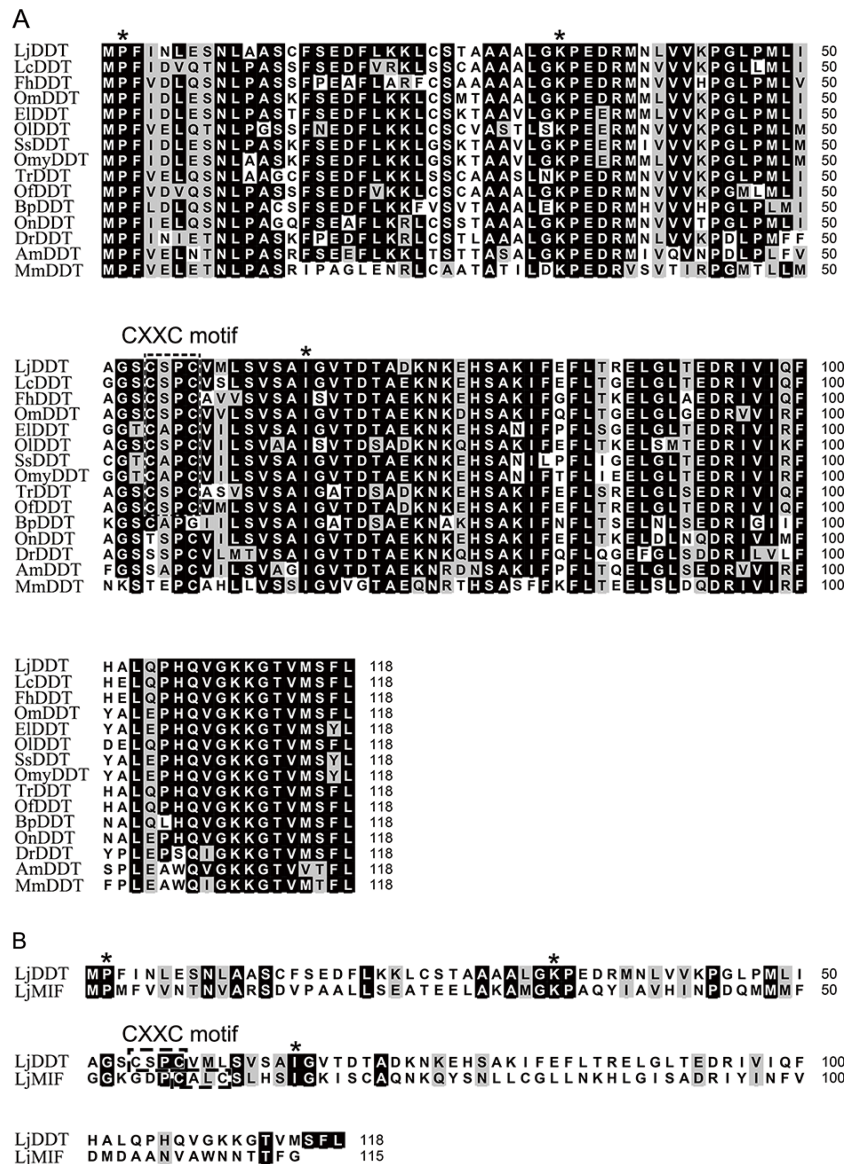
RESULTS

Sequence analysis of LjDDT

The cDNA sequence of LjDDT, 962 nucleotides (nts) in length, was deposited in the GenBank Data Library under accession No. MH988689. The sequence contained a large ORF of 357 nts, which encoded a 118 amino acid (aa) polypeptide with a calculated molecular weight (MW) of 12.7 kDa and a theoretical isoelectric point (pI) of 6.81. Sequence analysis revealed that LjDDT had no signal peptide (Figure 1A) and may be secreted through a non-classical mode (Supplementary Figure S1). Multiple alignments revealed that LjDDT had characteristic features of known DDT proteins. LjDDT contained the "CXXC" motif at aa position 54–57 and three conserved active site residues, Pro2, Lys33, and Ile65 (Figure 1A). LjDDT showed a similar structure to that of LjMIF (Figure 1B).

Sequence comparison revealed that *LjDDT* shared the highest nucleotide identity (90.76%) with rock bream *DDT*. Phylogenetic tree analysis showed that teleost fish *DDT*s grouped together to form a distinct subcluster closely related to the subcluster of higher vertebrate *DDT*s; LjDDT was most closely related to the rock bream homolog (Figure 2; Supplementary Figure S2); the DDT and MIF clusters were distantly related (Supplementary Figure S2).

Analysis of *LjDDT* mRNA expression in healthy and *V. harveyi*-infected Japanese sea bass

The mRNA expression levels of *LjDDT* in the tissues of healthy and *V. harveyi*-infected Japanese sea bass were investigated by qRT-PCR. In healthy fish, the *LjDDT* transcript was detected in all tested tissues, including the liver, spleen, trunk kidney, gill, intestine, brain, head kidney, heart, skin, and muscle, with the highest level detected in the liver, followed by the spleen and trunk kidney (Figure 3A). Upon *V. harveyi* infection, *LjDDT* transcripts were substantially down-regulated at 12 hpi or later in the liver, at 12 and 24 hpi in the head kidney, and at 6 hpi or later in the spleen (Figure 3B–D). The
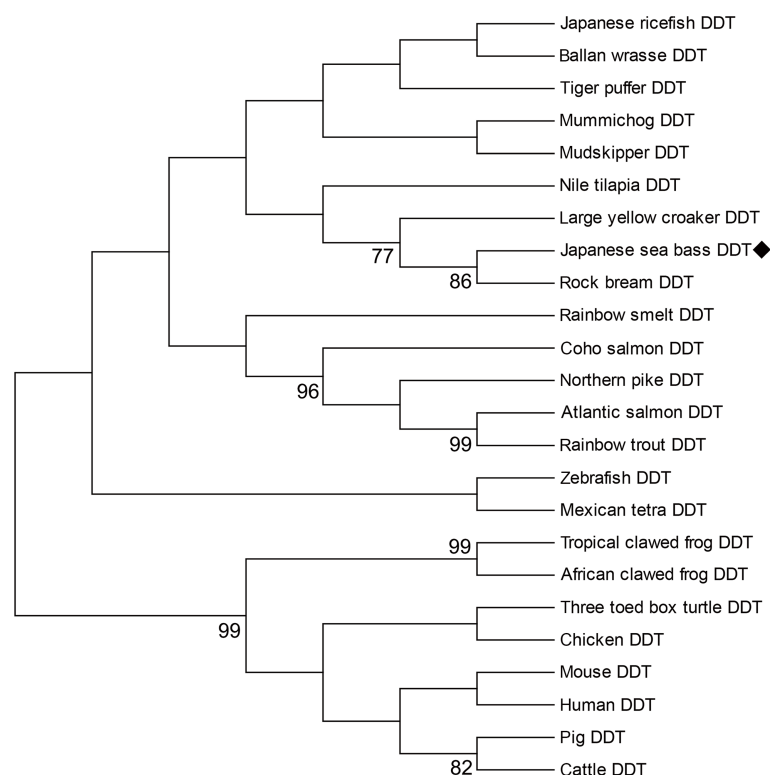
Figure 1 Multiple alignments of amino acid sequences of LjDDT with other DDT homologs (A) or LjMIF (B)

Threshold for shading was >60%, with similar residues shaded gray and identical residues shaded black. LjDDT: Japanese sea bass DDT; LcDDT: Large yellow croaker DDT; FhDDT: Mummichog DDT; OmDDT: Rainbow smelt DDT; ElDDT: Northern pike DDT; OlDDT: Japanese ricefish DDT; SsDDT: Atlantic salmon DDT; OmyDDT: Rainbow trout DDT; TrDDT: Tiger puffer DDT; OfDDT: Rock bream DDT; BpDDT: Mudskipper DDT; OnDDT: Nile tilapia DDT; DrDDT: Zebrafish DDT; AmDDT: Mexican tetra DDT; MmDDT: Mouse DDT; LjMIF: Japanese sea bass MIF. GenBank accession Nos. of sequences used are listed in Supplementary Table S1. Active site residues Pro, Lys, and Ile are marked with"*". "CXXC"motif is denoted with dotted box.

most significant *LjDDT* down-regulation was observed in the spleen at 36 hpi (0.31-fold) (Figure 3C).

Preparation of rLjDDT and corresponding antibody

After induction by IPTG, the recombinant Japanese sea bass DDT (rLjDDT) was overexpressed in *E. coli* BL21 (DE3). The MW of rLjDDT obtained from SDS-PAGE analysis was approximately 15 kDa, similar to the MW estimated from the sequence (12.7 kDa LjDDT plus 2.2 kDa His-tag) (Figure 4A).

Western blot analysis revealed that the MW of native LjDDT in the serum and liver of Japanese sea bass was approximately 13 kDa, similar to the MW calculated from the sequence (12.7 kDa) (Figure 4B).

Effect of rLjDDT and anti-rLjDDT on survival rate of *V. harveyi*-infected Japanese sea bass

The 9 d survival rate experiment investigated the effects of rLjDDT and anti-rLjDDT on *V. harveyi*-infected Japanese sea

Figure 2 Phylogenetic tree of *DDT* nucleotide using neighbor-joining method (1 000 bootstrap replicates; maximum composite likelihood model) in MEGA v7

Site of Japanese sea bass *DDT* is marked with"◆". Values at forks indicate percentage of trees in which this grouping occurred after bootstrapping (1 000 replicates; shown only when >60%). Scale bar shows number of substitutions per base. GenBank accession Nos. of sequences used are listed in Supplementary Table S1.
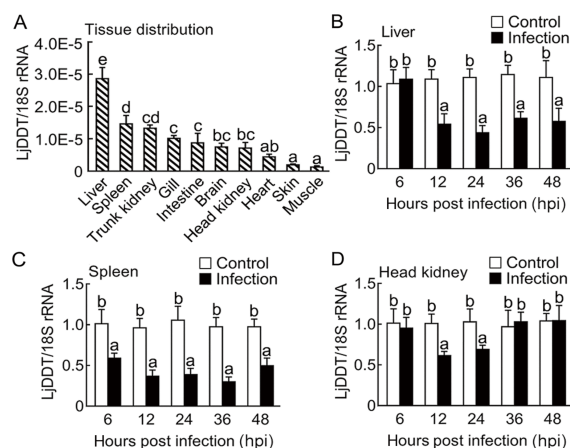

Figure 3 mRNA expression analysis of *LjDDT* in tissues of healthy (A) and *V. harveyi*-infected Japanese sea bass (B–D)

A: *LjDDT* mRNA expression level relative to that of *Lj18S* rRNA, calculated using $2^{-\Delta CT}$ method. B–D : Tissues were collected at different time points after bacterial infection. *LjDDT* mRNA expression levels relative to that of *Lj18S* rRNA were calculated using $2^{-\Delta\Delta CT}$ method. Data are expressed as mean±*SEM* of results from four fish. Values denoted by different letters are significantly different when compared by ANOVA (*P*<0.05).

Figure 4 Prokaryotic expression and Western blot analysis of LjDDT

A: 12% SDS-PAGE analysis of bacterial lysates and purified rLjDDT. Lane M: protein marker; Lane 1: pET-28a-LjDDT/BL21 before IPTG induction; Lane 2: pET-28a-LjDDT/BL21 after IPTG induction; Lane 3: Purified rLjDDT. B: Western blot analysis of rLjDDT and native LjDDT in liver of Japanese sea bass. Lane 4: pET28a-LjDDT/BL21 before IPTG induction, negative control; Lane 5: Purified rLjDDT; Lane 6: Japanese sea bass serum; Lane 7: Japanese sea bass liver lysates.

bass. Compared with the IsoIgG-treated group, fish administered with 10 µg/g or 100 µg/g anti-rLjDDT achieved a survival rate of 20% and 43.3%, respectively, at 9 d post infection (dpi), but only 100 µg/g anti-rLjDDT showed statistical significance (Figure 5). The administration of 100 µg/g rLjDDT accelerated the death of *V. harveyi*-infected fish, and all fish died at 7 dpi (Figure 5). Fish in the other five groups all died at 9 dpi (Figure 5).

In vitro chemotaxis assay of rLjDDT on different cells

In vitro transwell cell migration assay was conducted to test the chemotactic activity of rLjDDT and rLjMIF on MO/MΦ, lymphocytes, and neutrophils isolated from Japanese sea bass peripheral blood. Results showed that rLjDDT promoted the migration of MO/MΦ and lymphocytes in a dose-dependent manner (Figure 6A, B), but had no effect on the

Figure 5 Effect of LjDDT on survival rate of *V. harveyi*-infected Japanese sea bass

Fish were ip-injected with equal volumes of rLjDDT, IsoIgG, or anti-rLjDDT, respectively, 30 min after *V. harveyi* infection (1×10^4 CFU/fish) or 1 h before *V. harveyi* infection (1×10^4 CFU/fish). Control group received an equal volume of PBS. Fish mortality was monitored daily for 9 d. *n*=30.

migration of neutrophils (Figure 6C). Migration of MO/MΦ and lymphocytes was also inhibited by rLjMIF in a dose-dependent manner (Figure 6A, B), but had no effect on the migration of neutrophils (Figure 6C). The administration of rLjDDT combined with equivalent rLjMIF showed no significant effect on cell migration compared with the negative control (Figure 6A–C).

In vivo chemotaxis assay of rLjDDT on different cells

The numbers of migrated MO/MΦ, lymphocytes, and neutrophils in the abdominal cavity of Japanese sea bass were investigated 24 h after administration of rLjDDT and rLjMIF. rLjDDT administration induced a substantial increase in MO/MΦ (10.0 µg/g) and lymphocyte (1.0 or 10.0 µg/g) numbers in the abdominal cavity of Japanese sea bass compared with the control; no obvious change in neutrophil number was observed (Figure 7A–C). rLjMIF administration had no significant effect on MO/MΦ, lymphocyte, or neutrophil numbers in the abdominal cavity of Japanese sea bass compared with the negative control (Figure 7A–C). Only the administration of 10.0 µg/g rLjDDT+rLjMIF combined induced a substantial increase in MO/MΦ numbers in the abdominal cavity of Japanese sea bass compared with the negative control (Figure 7A–C).

In vitro effect of rLjDDT on migration of LPS- or cAMP-stimulated MO/MΦ

MO/MΦ polarization plays an important role in modulating proinflammatory responses in fish (Lu & Chen, 2019). The *in vitro* effect of rLjDDT on the migration of LPS- or cAMP-stimulated MO/MΦ was also determined. LPS- or cAMP-stimulation induced M1 and M2 polarization of Japanese sea bass MO/MΦ, respectively, with the up-regulation of iNOS (M1) and arginase activity (M2) (Figure 8A, B). rLjDDT promoted the migration of LPS-stimulated MO/MΦ (12.4% cells for 1.0 µg/mL rLjDDT, 15.3% cells for 10.0 µg/mL rLjDDT), whereas the random migration of LPS-stimulated MO/MΦ was 5.9% (Figure 8C). However, rLjDDT had no substantial effect on the migration of cAMP-stimulated MO/MΦ (Figure 8D).
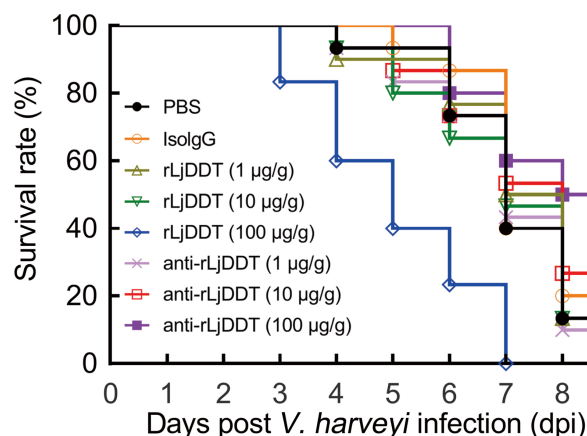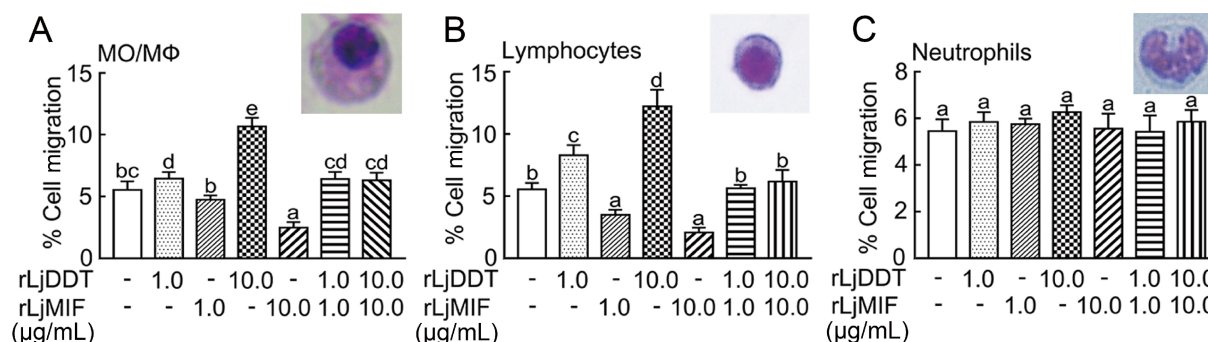

Figure 6 *In vitro* effect of rLjDDT and rLjMIF on migration of MO/MΦ (A), lymphocytes (B), and neutrophils (C) at different concentrations (0, 1.0, and 10.0 µg/mL, respectively)

Cells were counted under a light microscope at 400× magnification. Data are expressed as mean±*SEM*. *n*=4. Values denoted by different letters are significantly different when compared by ANOVA (*P*< 0.05).
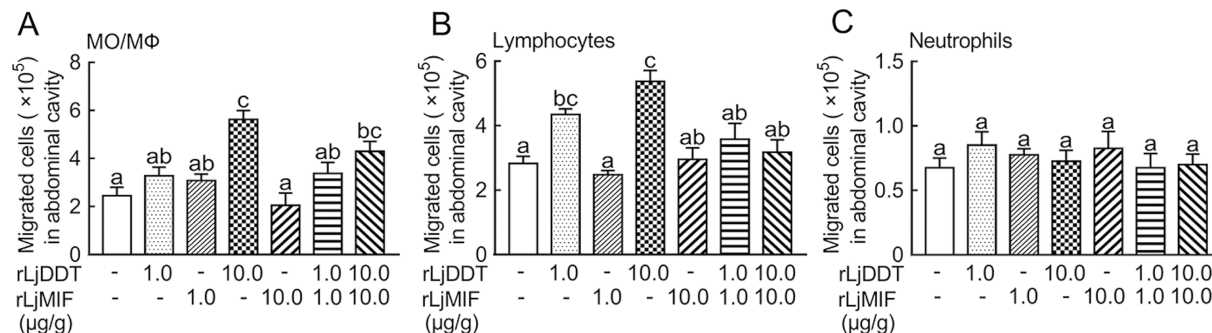
Figure 7 *In vivo* effect of rLjDDT and rLjMIF administration on MO/MΦ (A), lymphocyte (B), and neutrophil (C) numbers in abdominal cavity of Japanese sea bass at different concentrations (0, 1.0, and 10.0 µg/g respectively)

Cells were counted under a light microscope at 400× magnification 24 h after administration of rLjDDT and rLjMIF. Data are expressed as mean± SEM. *n*=4. Values denoted by different letters are significantly different when compared by ANOVA (*P*<0.05).

Effects of *LjCD74* and *LjCXCR4* knockdown on rLjDDT and rLjMIF-induced migration of MO/MΦ

As CD74 and CXCR4 are considered receptors of DDT in mammals (Fagone et al., 2018; Klasen et al., 2014), we determined whether *LjCD74* and *LjCXCR4* knockdown influenced rLjDDT-induced migration of MO/MΦ. We first used RNAi to knock down the expression of *LjCD74* and *LjCXCR4* in Japanese sea bass MO/MΦ. When MO/MΦ were transfected with LjCD74si or LjCXCR4si, the mRNA expression of *LjCD74* and *LjCXCR4* decreased to 23.38%± 8.05% and 21.79%±4.44%, respectively, of the normal control

at 48 h (Figure 9A, B), suggesting that *LjCD74* and *LjCXCR4* were effectively knocked down by LjCD74si and LjCXCR4si, respectively. The transfection of MsiRNA had no obvious effect on *LjCD74* or *LjCXCR4* expression (Figure 9A, B). We next used LjCD74si and LjCXCR4si to explore whether *LjCD74* and *LjCXCR4* mediated the effect of LjDDT on MO/MΦ migration. After transfection with MsiRNA, 11.02% and 2.04% of MO/M Φ migrated to the lower chambers containing 10.0 µg/mL rLjDDT and rLjMIF, respectively (Figure 9C). Only 6.13% of MO/MΦ migrated to the lower chambers
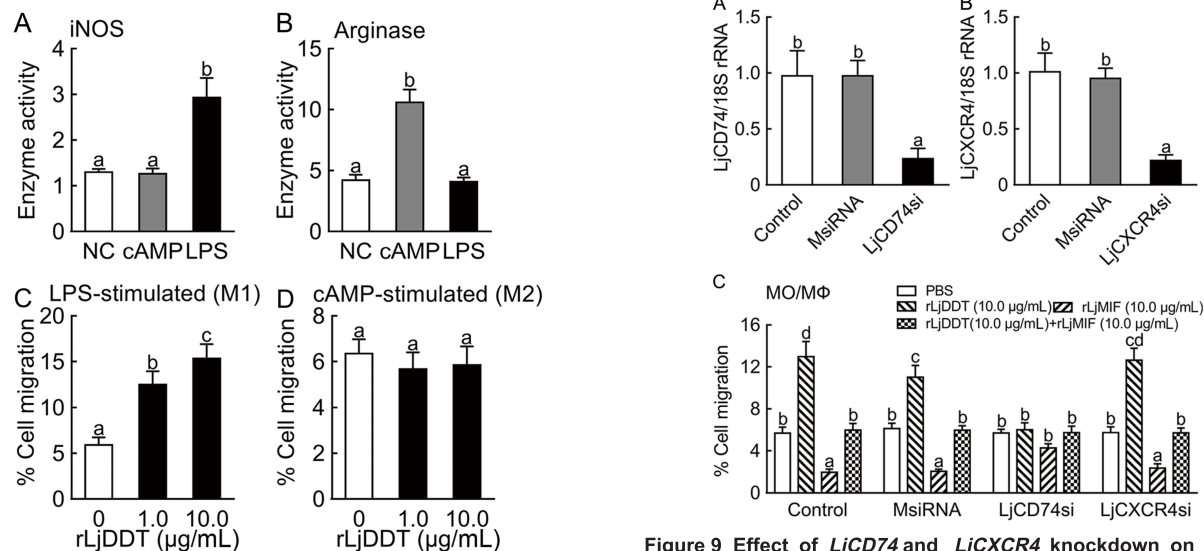

Figure 8 Effect of rLjDDT on migration of polarized Japanese sea bass MO/MΦ

LPS and cAMP were used to induce M1 and M2 polarization of MO/MΦ, respectively. Activities of iNOS (A) and arginase (B) were determined. After incubation with rLjDDT for 4 h, migration percentage of LPS- (C) or cAMP- (D) stimulated MO/MΦ was determined. Non-stimulated resting MO/MΦ were used as negative control (NC). Data are expressed as mean±SEM. *n*=4; Values denoted by different letters are significantly different when compared by ANOVA (*P*<0.05).

Figure 9 Effect of *LjCD74* and *LjCXCR4* knockdown on rLjDDT and rLjMIF-induced migration of MO/MΦ, respectively

Histogram displays effect of *LjCD74* (A) and *LjCXCR4* (B) siRNA transfection on knockdown of MO/MΦ *LjCD74* and *LjCXCR4* mRNA expression by RT-qPCR analysis. C: Migration percentage of Japanese sea bass MO/MΦ was examined in a Transwell chamber in presence or absence of 10.0 µg/mL rLjDDT, rLjMIF, or rLjDDT+rLjMIF combined. Each bar represents mean±SEM, *n*=4. Values denoted by different letters are significantly different when compared by ANOVA (*P*<0.05).

without rLjDDT or rLjMIF (Figure 9C). After knockdown of *LjCD74*, approximately 6.01% and 4.27% of the MO/MΦ migrated to the lower chambers containing 10.0 µg/mL rLjDDT and rLjMIF, respectively, compared to 5.74% in the group treated with 10.0 µg/mL rLjDDT+rLjMIF combined (Figure 9C). After knockdown of *LjCXCR4* , 12.63% and 2.36% of the MO/MΦ migrated to the lower chambers containing 10.0 µg/mL rLjDDT and rLjMIF, respectively, compared to 5.71% in the group treated with 10.0 µg/mL rLjDDT+rLjMIF combined (Figure 9C).

DISCUSSION

As a second member of the MIF superfamily, DDT is involved in various pathological roles in inflammatory, autoimmune, and chronic respiratory mammalian diseases (Günther et al., 2019; Jankauskas et al., 2019; Sinitski et al., 2019). It is clear that DDT and MIF are pleiotropic cytokines in mammals, which not only share an overlapping spectrum of activities, but also distinct functions (Benedek et al., 2017; Furukawa et al., 2016; Tilstam et al., 2017; Vincent et al., 2018). In the present study, we identified one gene-encoding DDT homolog (LjDDT) in Japanese sea bass. Without determination of complete genomic sequences, we cannot know whether the Japanese sea bass has two or more *DDT*s. Sequence comparison and phylogenetic tree analysis revealed that LjDDT was highly conserved and most closely related to the rock bream homolog. LjDDT also shared 29.2% aa identity with LjMIF, and their canonical N-terminal proline residues and enzyme activity-related sites were highly conserved. LjDDT also lacked an N-terminal signal peptide or an internal secretory sequence, thus it may be released from cells via a noncanonical protein secretion pathway, like MIF (Merk et al., 2009). The sequence similarity between LjDDT and LjMIF suggests that their biological functions may be highly correlated.

Assessment of gene expression profiles should help clarify the functionality of the *DDT* gene. In mammals, earlier studies revealed that *DDT* is constitutively expressed in rat and human tissues, with the highest level found in the liver (Nishihira et al., 1998; Zhang et al., 1995). Further studies have found that the expression of *DDT* is altered under pathological conditions. For instance, both serum protein and mRNA expression levels of *DDT* are significantly higher in burn patients compared to healthy individuals (Kim et al., 2016); whereas *DDT* mRNA expression is down-regulated in inflammatory adipose tissue in patients with wound healing disorders (Kim et al., 2017). DDT homologs in fish have also been found in many species, although their expression profiles have been rarely studied (Oh et al., 2013; Shen et al., 2012). In zebrafish, *DDT* transcripts have been detected in whole embryos throughout embryogenesis (Shen et al., 2012). In rock bream, *DDT* transcripts have been ubiquitously detected in all tested tissues, with the highest expression found in the liver, followed by blood, heart, and kidneys (Oh et al., 2013). In the present study, *LjDDT* was found to be constitutively expressed in all tested tissues of healthy Japanese sea bass,

with the highest expression in the liver, consistent with previously reported results in other animals (Nishihira et al., 1998; Oh et al., 2013; Zhang et al., 1995). After *V. harveyi* infection, *LjDDT* mRNA expression was significantly down-regulated in the three tested immune tissues, i. e., liver, spleen, and head kidney. The expression profile of *LjDDT* was different to that of *LjMIF*, as determined in our previous work (Xu et al., 2019). This suggests that *LjDDT* may be functionally different from *LjMIF* , which coincides with previously described human results (Kim et al., 2017). However, the highest expression levels of *LjDDT* and *LjMIF* were all found in the liver (Xu et al., 2019), indicating that *DDT* expression is closely associated with the immune response of Japanese sea bass against *V. harveyi*.

Survival rates can intuitively reflect the degree of damage to healthy organisms caused by pathogenic infections (Fernández et al., 2018). In mammals, DDT is associated with host immunity in relation to inflammatory responses and disease severity (Merk et al., 2011; Pohl et al., 2017; Valiño-Rivas et al., 2018). For example, the administration of a specific anti-DDT antibody protects mice from lethal endotoxemia (Merk et al., 2011). In this study, we found that the administration of 10 µg/g and 100 µg/g anti-rLjDDT reduced mortality in *V. harveyi*-infected fish, but only the 100 µg/g anti-rLjDDT treatment showed statistical significance. On the other hand, the administration of 100 µg/g rLjDDT accelerated death in *V. harveyi*-infected fish. Our results coincide well with those reported previously in mice (Merk et al., 2011), suggesting that LjDDT plays a role in Japanese sea bass immunity.

Immune cell migration is a key component of many pathological processes, such as inflammation and cancer metastasis, which makes it an exciting and crucial field of study (Luster et al., 2005). MIF has a well-known chemokine-like function involving the trafficking and recruitment of macrophages and lymphocytes in vertebrates (Abe et al., 2001; Bernhagen et al., 2007; Jin et al., 2007; Schober et al., 2008; Xu et al., 2019), but there is little information on the chemotactic activity of DDT (Kim et al., 2017; Merk et al., 2011; Pasupuleti et al., 2014). DDT inhibits chemotaxis of human peripheral blood monocytes to monocyte chemoattractant protein-1 (MCP-1) (Merk et al., 2011) and functionally cooperates with MIF in promoting endothelial cell migration in the development of renal carcinoma (Pasupuleti et al., 2014). Injection of LPS combined with MIF can lead to higher peritoneal macrophage accumulation in mouse epididymal fat pads compared with the LPS-group, but with no such effect for LPS combined with DDT (Kim et al., 2017). In this study, we found that rLjDDT induced the migration of MO/MΦ and lymphocytes both *in vitro* and *in vivo*. In our previous study, we found that rLjMIF can inhibit the migration of MO/MΦ and lymphocytes (Xu et al., 2019). The opposite effect of rLjDDT and rLjMIF on immune cells suggests that they antagonistically regulate MO/MΦ and lymphocyte trafficking. MO/MΦ polarization plays an important role in modulating proinflammatory responses in fish (Lu & Chen, 2019). In teleosts, LPS from gram-negative bacteria can induce M1 polarization and cAMP can induce M2 polarization (Joerink et

al., 2006). In this study, we investigated the chemotactic activity of rLjDDT on polarized Japanese sea bass MO/MΦ. We found that rLjDDT exhibited chemotactic activity for LPS-stimulated M1-type MO/MΦ, but not for cAMP-stimulated M2-type MO/MΦ. As M1 macrophages are proinflammatory (Wang et al., 2019), our results suggest that LjDDT may be involved in the proinflammatory responses of Japanese sea bass.

The receptor mechanisms by which MIF activates target cells have long been unclear. MIF is known to not only interact with CD74, but also bind to CXCR2 and CXCR4 (Klasen et al., 2014; Presti et al., 2018; Schwartz et al., 2009; Soppert et al., 2018; Weber et al., 2008). MIF participates in the recruitment of many cell types via CXCR4 (Pawig et al., 2015). CD74 can form functional complexes with CXCR4 to mediate MIF-specific signaling (Schwartz et al., 2009). DDT also binds to CD74 with high affinity (Merk et al., 2011), but lacks the essential motif for binding to CXCR2 (Merk et al., 2012). In this study, compared with normal and MsiRNA groups, the knockdown of *LjCD74* expression in MO/MΦ significantly decreased the rLjDDT-enhanced migration of MO/MΦ, and relieved the rLjMIF-inhibited migration of MO/MΦ. The knockdown of *LjCXCR4* had no significant influence on rLjDDT-enhanced or rLjMIF-inhibited migration of MO/MΦ. The combination treatment of rLjDDT+rLjMIF had no significant effect on the migration of normal, MsiRNA, LjCD74si, or LjCXCR4si-treated MO/MΦ. This suggests that LjDDT and LjMIF have an antagonistic effect on the migration of Japanese sea bass MO/MΦ through the mediation of *LjCD74*, but not of *LjCXCR4*.

In conclusion, we characterized a DDT gene from Japanese sea bass. Upon *V. harveyi* infection, the *LjDDT* expression profiles were significantly altered in immune tissues. Antibody neutralization of LjDDT had protective effects on the survival rate of *V. harveyi*-infected Japanese sea bass. *In vivo* and *in vitro* studies revealed that LjDDT participates in the immune response by mediating the trafficking of lymphocytes and resting and M1-type MO/MΦ. After knocking down the expression of *LjCD74*, the chemotaxis of rLjDDT on MO/MΦ decreased significantly. Our present work investigated the primary role of LjDDT in Japanese sea bass immune responses. Further studies on the precise chemotactic mechanism of LjDDT and LjMIF release in response to pathogenic infections should provide insight into their immunological functions.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

J.C. and M.Y.L. drafted the experiments; F.X. performed the experiments. F. X. and J. C. analyzed the data and wrote the paper. All authors read and approved the final version of the manuscript.

REFERENCES

Abe R, Peng T, Sailors J, Bucala R, Metz CN. 2001. Regulation of the CTL response by macrophage migration inhibitory factor. *The Journal of Immunology*, **166**(2): 747−753.

Benedek G, Meza-Romero R, Jordan K, Zhang Y, Nguyen H, Kent G, Li J, Siu E, Frazer J, Piecychna M, Du X, Sreih A, Leng L, Wiedrick J, Caillier SJ, Offner H, Oksenberg JR, Yadav V, Bourdette D, Bucala R, Vandenbark AA. 2017. MIF and D-DT are potential disease severity modifiers in male MS subjects. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(40): E8421−E8429.

Bernhagen J, Krohn R, Lue H, Gregory JL, Zernecke A, Koenen RR, Dewor M, Georgiev I, Schober A, Leng L, Kooistra T, Fingerle-Rowson G, Ghezzi P, Kleemann R, McColl SR, Bucala R, Hickey MJ, Weber C. 2007. MIF is a noncognate ligand of CXC chemokine receptors in inflammatory and atherogenic cell recruitment. *Nature Medicine*, **13**(5): 587−596.

Bloom BR, Bennett B. 1966. Mechanism of a reaction in vitro associated with delayed-type hypersensitivity. *Science*, **153**(3731): 80−82.

Chen F, Lu XJ, Nie L, Ning YJ, Chen J. 2018. Molecular characterization of a CC motif chemokine 19-like gene in ayu (*Plecoglossus altivelis*) and its role in leukocyte trafficking. *Fish and Shellfish Immunology*, **72**: 301−308.

Chen J, Chen Q, Lu XJ, Li CH. 2014. LECT2 improves the outcomes in ayu with *Vibrio anguillarum* infection via monocytes/macrophages. *Fish and Shellfish Immunology*, **41**(2): 586−592.

Chen K, Shi YH, Chen J, Li MY. 2019. A soluble FcγR homolog inhibits IgM antibody production in ayu spleen cells. *Zoological Research*, **2019, 40**(5): 404−415.

Coleman AM, Rendon BE, Zhao M, Qian MW, Bucala R, Xin D, Mitchell RA. 2008. Cooperative regulation of non-small cell lung carcinoma angiogenic potential by macrophage migration inhibitory factor and its homolog, D-dopachrome tautomerase. *The Journal of Immunology*, **181**(4): 2330−2337.

David JR. 1966. Delayed hypersensitivity *in vitro*: its mediation by cell-free substances formed by lymphoid cell-antigen interaction. *Proceedings of the National Academy of Sciences of the United States of America*, **56**(1): 72−77.

Esumi N, Budarf M, Ciccarelli L, Sellinger B, Kozak CA, Wistow G. 1998. Conserved gene structure and genomic linkage for D-dopachrome tautomerase (DDT) and MIF. *Mammalian Genome*, **9**(9): 753−757.

Fagone P, Mazzon E, Cavalli E, Bramanti A, Petralia MC, Mangano K, Al-Abed Y, Bramati P, Nicoletti F. 2018. Contribution of the macrophage migration inhibitory factor superfamily of cytokines in the pathogenesis of preclinical and human multiple sclerosis: in silico and in vivo evidences. *Journal of Neuroimmunology*, **322**: 46−56.

Fernández J, Acevedo J, Wiest R, Gustot T, Amoros A, Deulofeu C, Reverter E, Martínez J, Saliba F, Jalan R, Welzel T, Pavesi M, Hernández-Tejero M, Ginès P, Arroyo V. 2018. Bacterial and fungal infections in acute-on-chronic liver failure: prevalence, characteristics and impact on prognosis. *Gut*, **67**(10): 1870−1880.

Furukawa R, Tamaki K, Kaneko H. 2016. Two macrophage migration inhibitory factors regulate starfish larval immune cell chemotaxis. *Immunology and Cell Biology*, **94**(4): 315−321.

Günther S, Fagone P, Jalce G, Atanasov AG, Guignabert C, Nicoletti F. 2019. Role of MIF and D-DT in immune-inflammatory, autoimmune, and chronic respiratory diseases: from pathogenic factors to therapeutic targets.

Drug Discovery Today, **24**(2): 428−439.

Guo D, Guo J, Yao J, Jiang K, Hu J, Wang B, Liu H, Sun W, Jiang X. 2016. D-dopachrome tautomerase is over-expressed in pancreatic ductal adenocarcinoma and acts cooperatively with macrophage migration inhibitory factor to promote cancer growth. *International Journal of Cancer*, **139**(9): 2056−2067.

Jankauskas SS, Wong DWL, Bucala R, Djudjaj S, Boor P. 2019. Evolving complexity of MIF signaling. *Cellular Signalling*, **57**: 76−88.

Jin HJ, Xiang LX, Shao JZ. 2007. Molecular cloning and identification of macrophage migration inhibitory factor (MIF) in teleost fish. *Developmental and Comparative Immunology*, **31**(11): 1131−1144.

Joerink M, Ribeiro CMS, Stet RJM, Hermsen T, Savelkoul HFJ, Wiegertjes GF. 2006. Head kidney-derived macrophages of common carp (*Cyprinus carpio* L.) *show plasticity and functional polarization upon differential stimulation. The Journal of Immunology*, **177**(1): 61−69.

Kim BS, Stoppe C, Grieb G, Leng L, Sauler M, Assis D, Simons D, Boecker AH, Schulte W, Piecychna M, Hager S, Bernhagen J, Pallua N, Bucala R. 2016. The clinical significance of the MIF homolog d-dopachrome tautomerase (MIF-2) and its circulating receptor (sCD74) in burn. *Burns*, **42**(6): 1265−1276.

Kim BS, Tilstam PV, Hwang SS, Simons D, Schulte W, Leng L, Sauler M, Ganse B, Averdunk L, Kopp R, Stoppe C, Bernhagen J, Pallua N, Bucala R. 2017. D-dopachrome tautomerase in adipose tissue inflammation and wound repair. *Journal of Cellular and Molecular Medicine*, **21**(1): 35−45.

Klasen C, Ohl K, Sternkopf M, Shachar I, Schmitz C, Heussen N, Hobeika E, Levit-Zerdoun E, Tenbrock K, Reth M, Bernhagen J, El Bounkari O. 2014. MIF promotes B cell chemotaxis through the receptors CXCR4 and CD74 and ZAP-70 signaling. *The Journal of Immunology*, **192**(11): 5273−5284.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870−1874.

Leng L, Bucala R. 2006. Insight into the biology of macrophage migration inhibitory factor (MIF) revealed by the cloning of its cell surface receptor. *Cell Research*, **16**(2): 162−168.

Liu H, Lu XJ, Chen J. 2018. Full-length and a smaller globular fragment of adiponectin have opposite roles in regulating monocyte/macrophage functions in ayu, *Plecoglossus altivelis*. *Fish and Shellfish Immunology*, **82**: 319−329.

Lu XJ, Chen J. 2019. Specific function and modulation of teleost monocytes/ macrophages: polarization and phagocytosis. *Zoological Research*, **40**(3): 146−150.

Luster AD, Alon R, von Andrian UH. 2005. Immune cell migration in inflammation: present and future therapeutic targets. *Nature Immunology*, **6**(12): 1182−1190.

Ma Y, Su KN, Pfau D, Rao VS, Wu X, Hu X, Leng L, Du X, Piecychna M, Bedi K, Campbell SG, Eichmann A, Testani JM, Margulies KB, Bucala R, Young LH. 2019. Cardiomyocyte d-dopachrome tautomerase protects against heart failure. *The Journal of Clinical Investigation Insight*, **4**(17): 128900.

Merk M, Baugh J, Zierow S, Leng L, Pal U, Lee SJ, Ebert AD, Mizue Y, Trent JO, Mitchell R, Nickel W, Kavathas PB, Bernhagen J, Bucala R. 2009. The golgi-associated protein p115 mediates the secretion of macrophage migration inhibitory factor. *The Journal of Immunology*, **182**(11):

6896−6906.

Merk M, Mitchell RA, Endres S, Bucala R. 2012. D-dopachrome tautomerase (D-DT or MIF-2): doubling the MIF cytokine family. *Cytokine*, **59**(1): 10−17.

Merk M, Zierow S, Leng L, Das R, Du X, Schulte W, Fan J, Lue H, Chen Y, Xiong H, Chagnon F, Bernhagen J, Lolis E, Mor G, Lesur O, Bucala R. 2011. The D-dopachrome tautomerase (DDT) gene product is a cytokine and functional homolog of macrophage migration inhibitory factor (MIF). *Proceedings of the National Academy of Sciences of the United States of America*, **108**(34): E577−E585.

Nishihira J, Fujinaga M, Kuriyama T, Suzuki M, Sugimoto H, Nakagawa A, Tanaka I, Sakai M. 1998. Molecular cloning of human D-dopachrome tautomerase cDNA: N-terminal proline is essential for enzyme activation. *Biochemical and Biophysical Research Communications*, **243**(2): 538−544.

Odh G, Hindemith A, Rosengren AM, Rosengren E, Rorsman H. 1993. Isolation of a new tautomerase monitored by the conversion of D-dopachrome to 5,6-dihydroxyindole. *Biochemical and Biophysical Research Communications*, **197**(2): 619−624.

Oh M, Kasthuri SR, Wan Q, Bathige SDNK, Whang I, Lim BS, Jung HB, Oh MJ, Jung SJ, Kim SY, Lee J. 2013. Characterization of MIF family proteins: MIF and DDT from rock bream, *Oplegnathus fasciatus*. *Fish and Shellfish Immunology*, **35**(2): 458−468.

Pasupuleti V, Du W, Gupta Y, Yeh IJ, Montano M, Magi-Galuzzi C, Welford SM. 2014. Dysregulated D-dopachrome tautomerase, a hypoxia-inducible factor-dependent gene, cooperates with macrophage migration inhibitory factor in renal tumorigenesis. *Journal of Biological Chemistry*, **289**(6): 3713−3723.

Pawig L, Klasen C, Weber C, Bernhagen J, Noels H. 2015. Diversity and inter-connections in the CXCR4 chemokine receptor/ligand family: molecular perspectives. *Frontiers in Immunology*, **6**: 429.

Pohl J, Hendgen-Cotta UB, Stock P, Luedike P, Rassaf T. 2017. Elevated MIF-2 levels predict mortality in critically ill patients. *Journal of Critical Care*, **40**: 52−57.

Presti M, Mazzon E, Basile MS, Petralia MC, Bramanti A, Colletti G, Bramanti P, Nicoletti F, Fagone P. 2018. Overexpression of macrophage migration inhibitory factor and functionally-related genes, D-DT, CD74, CD44, CXCR2 and CXCR4, in glioblastoma. *Oncology Letters*, **16**(3): 2881−2886.

Rajasekaran D, Gröning S, Schmitz C, Zierow S, Drucker N, Bakou M, Kohl K, Mertens A, Lue H, Weber C, Xiao A, Luker G, Kapurniotu A, Lolis E, Bernhagen J. 2016. Macrophage migration inhibitory factor-CXCR4 receptor interactions: evidence for partial allosteric agonism in comparison with CXCL12 chemokine. *Journal of Biological Chemistry*, **291**(30): 15881−15895.

Ren Y, Liu SF, Nie L, Cai SY, Chen J. 2019. Involvement of ayu NOD2 in NF-κB and MAPK signaling pathways: insights into functional conservation of NOD2 in antibacterial innate immunity. *Zoological Research*, **40**(2): 77−88.

Rijvers L, Melief MJ, van der Vuurst de Vries RM, Stéphant M, van Langelaar J, Wierenga-Wolf AF, Hogervorst JM, Geurts-Moespot AJ, Sweep FCGJ, Hintzen RQ, van Luijn MM. 2018. The macrophage migration inhibitory factor pathway in human B cells is tightly controlled and dysregulated in multiple sclerosis. *European Journal of Immunology*, **48**(11): 1861−1871.

Schober A, Bernhagen J, Weber C. 2008. Chemokine-like functions of MIF in atherosclerosis. *Journal of Molecular Medicine*, **86**(7): 761−770.

Schwartz V, Lue H, Kraemer S, Korbiel J, Krohn R, Ohl K, Bucala R, Weber C, Bernhagen J. 2009. A functional heteromeric MIF receptor formed by CD74 and CXCR4. *FEBS Letters*, **583**(17): 2749−2757.

Shen YC, Thompson DL, Kuah MK, Wong KL, Wu KL, Linn SA, Jewett EM, Shu-Chien AC, Barald KF. 2012. The cytokine macrophage migration inhibitory factor (MIF) acts as a neurotrophin in the developing inner ear of the zebrafish, *Danio rerio*. *Developmental Biology*, **363**(1): 84−94.

Sinitski D, Kontos C, Krammer C, Asare Y, Kapurniotu A, Bernhagen J. 2019. Macrophage migration inhibitory factor (MIF)-based therapeutic concepts in atherosclerosis and inflammation. *Journal of Thrombosis and Haemostasis*, **119**(4): 553−566.

Soppert J, Kraemer S, Beckers C, Averdunk L, Möllmann J, Denecke B, Goetzenich A, Marx G, Bernhagen J, Stoppe C. 2018. Soluble CD74 reroutes MIF/CXCR4/AKT-mediated survival of cardiac myofibroblasts to necroptosis. *Journal of the American Heart Association*, **7**(17): e009384.

Sugimoto H, Taniguchi M, Nakagawa A, Tanaka I, Suzuki M, Nishihira J. 1999. Crystal structure of human D-dopachrome tautomerase, a homologue of macrophage migration inhibitory factor, at1.54 Å resolution. *Biochemistry*, **38**(11): 3268−3279.

Tilstam PV, Qi D, Leng L, Young L, Bucala R. 2017. MIF family cytokines in cardiovascular diseases and prospects for precision-based therapeutics. *Expert Opinion on Therapeutic Targets*, **21**(7): 671−683.

Valiño-Rivas L, Cuarental L, Grana O, Bucala R, Leng L, Sanz A, Gomez G, Ortiz A, Sanchez-Niño MD. 2018. TWEAK increases CD74 expression and sensitizes to DDT proinflammatory actions in tubular cells. *PLoS One*, **13**(6): e0199391.

Vincent FB, Lin E, Sahhar J, Ngian GS, Kandane-Rathnayake R, Mende R, Hoi AY, Morand EF, Lang T, Harris J. 2018. Analysis of serum macrophage migration inhibitory factor and D-dopachrome tautomerase in systemic sclerosis. *Clinical & Translational Immunology*, **7**(12): e1042.

Wang LX, Zhang SX, Wu HJ, Rong XL, Guo J. 2019. M2b macrophage polarization and its roles in diseases. *Journal of Leukocyte Biology*, **106**(2): 345−358.

Wang Q, Wei Y, Zhang J. 2017. Combined knockdown of D-dopachrome tautomerase and migration inhibitory factor inhibits the proliferation, migration, and invasion in human cervical cancer. *International Journal of Gynecological Cancer*, **27**(4): 634−642.

Weber C, Kraemer S, Drechsler M, Lue H, Koenen RR, Kapurniotu A, Zernecke A, Bernhagen J. 2008. Structural determinants of MIF functions in CXCR2-mediated inflammatory and atherogenic leukocyte recruitment. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(42): 16278−16283.

Xu F, Shi YH, Chen J. 2019. Characterization and immunologic functions of the macrophage migration inhibitory factor from Japanese sea bass, *Lateolabrax japonicus*. *Fish and Shellfish Immunology*, **86**: 947−955.

Yu L, Li CH, Chen J. 2019. A novel CC chemokine ligand 2 like gene from ayu *Plecoglossus altivelis* is involved in the innate immune response against to *Vibrio anguillarum*. *Fish and Shellfish Immunology*, **87**: 886−896.

Zhang M, Åman P, Grubb A, Panagopoulos I, Hindemith A, Rosengren E, Rorsman H. 1995. Cloning and sequencing of a cDNA encoding rat D-dopachrome tautomerase. *FEBS Letters*, **373**(3): 203−206.

Zhou QJ, Wang L, Chen J, Wang RN, Shi YH, Li CH, Zhang DM, Yan XJ, Zhang YJ. 2014. Development and evaluation of a real-time fluorogenic loop-mediated isothermal amplification assay integrated on a microfluidic disc chip (on-chip LAMP) for rapid and simultaneous detection of ten pathogenic bacteria in aquatic animals. *Journal of Microbiological Methods*, **104**: 26−35.

Potential dual expansion of domesticated donkeys revealed by worldwide analysis on mitochondrial sequences

Xi-Yao Ma[1,#], Tiao Ning[2,3,#], Adeniyi C. Adeola[4,5,#], Jie Li[1,#], Ali Esmailizadeh[4,6], Jacqueline K. Lichoti[7], Bernard R. Agwanda[8], Jainagul Isakova[9], Almaz A. Aldashev[9], Shi-Fang Wu[4], He-Qun Liu[4], Najmudinov Tojiddin Abdulloevich[10], Manilova Elena Afanasevna[10], Khudoidodov Behruz Ibrohimovich[10], Rahamon Akinyele Moshood Adedokun[11], Sunday Charles Olaogun[11], Oscar J. Sanke[12], Godwin F. Mangbon[13], Xi Chen[14,15], Wei-Kang Yang[14,15], Zhe Wang[16,17], Min-Sheng Peng[4,5,18], Sheila C. Ommeh[19,*], Yan Li[1,*], Ya-Ping Zhang[1,4,18,20,*]

[1] *State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, School of Life Sciences, Yunnan University, Kunming, Yunnan 650091, China*

[2] *Agriculture College, Kunming University, Kunming, Yunnan 650214, China*

[3] *Engineering Research Center for Urban Modern Agriculture of Higher Education in Yunnan Province, Kunming University, Kunming, Yunnan 650214, China*

[4] *State Key Laboratory of Genetic Resources and Evolution, Yunnan Laboratory of Molecular Biology of Domestic Animals, Germplasm Bank of Wild Species, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

[5] *Sino-Africa Joint Research Center, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

[6] *Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman PB 76169-133, Iran*

[7] *State Department of Livestock, Ministry of Agriculture Livestock and Fisheries, Nairobi 00100, Kenya*

[8] *Department of Zoology, National Museums of Kenya, Nairobi 00100, Kenya*

[9] *Institute of Molecular Biology and Medicine, Bishkek 720040, Kyrgyzstan*

[10] *E.N. Pavlovsky Institute of Zoology and Parasitology, Academy of Sciences of Republic of Tajikistan, Dushanbe 734025, Tajikistan*

[11] *Department of Veterinary Medicine, University of Ibadan, Ibadan 200284, Nigeria*

[12] *Taraba State Ministry of Agriculture and Natural Resources, Jalingo 660221, Nigeria*

[13] *Division of Veterinary Office, Serti 663101, Nigeria*

[14] *Research Center for Ecology and Environment of Central Asia, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China*

[15] *Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China*

[16] *Department of Physiology, Hebei Medical University, Shijiazhuang, Hebei 050017, China*

[17] *Hebei Key Laboratory of Animal Science, Hebei Medical University, Shijiazhuang, Hebei 050017, China*

[18] *Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China*

[19] *Animal Biotechnology Group, Institute of Biotechnology Research, Jomo Kenyatta University of Agriculture and Technology, Nairobi 00200, Kenya*

[20] *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

ABSTRACT

Molecular studies on donkey mitochondrial sequences have clearly defined two distinct maternal lineages involved in domestication. However, domestication histories of these two lineages remain enigmatic. We therefore compared several population characteristics between these two lineages based on global sampling, which included 171 sequences obtained in this study (including Middle Asian, East Asian, and African samples) plus 536 published sequences (including European, Asian, and African samples). The two lineages were clearly separated from each other based on whole mitochondrial genomes and partial non-coding displacement loop (D-loop) sequences, respectively. The Clade I lineage experienced an increase in population size more than 8 000 years ago and shows a complex haplotype network. In contrast, the population size of the Clade II lineage has remained relatively constant, with a simpler haplotype network. Although the distribution of the two lineages was almost equal across the Eurasian mainland, they still presented discernible but complex geographic bias in most parts of Africa, which are known as their domestication sites. Donkeys from sub-Saharan Africa tended to descend from the Clade I lineage, whereas the Clade II lineage was dominant along the East and North coasts of Africa. Furthermore, the migration routes inferred from diversity decay suggested different expansion across China between the two lineages. Altogether, these differences indicated non-simultaneous domestication of the two lineages, which was possibly influenced by the response of pastoralists to the desertification of the Sahara and by the social expansion and trade of ancient humans in Northeast Africa, respectively.

Keywords: Donkey lineage; Domestication history; Population; Expansion

INTRODUCTION

Unlike other species with a similar historical function (e.g., horses), domestic donkeys are underrepresented in the scientific literature (Blench, 2000). Importantly, donkeys remain an essential means of transport in modern society for people living in mountain areas, deserts, and poorer regions of the world (Smith & Pearson, 2005; Starkey, 2000). Both molecular data and archaeological evidence strongly support an African origin for the domestic donkey (*Equus asinus*) (Beja-Pereira et al., 2004; Kimura et al., 2011; Rossel et al., 2008). This domestication, together with that of horses, is believed to have contributed to mobile pastoralism as well as the establishment of ancient overland trade routes and spread of ancient civilizations (Rosenbom et al., 2015). Previous studies on mitochondrial DNA (mtDNA) have revealed that modern donkeys can be clearly separated into two distinct clades: one clustered directly with the Nubian wild ass (*E. africanus africanus*), hereafter termed the Clade I lineage; another derived from a probably extinct wild ass close to the Somali wild ass (*E. africanus somaliensis*), hereafter termed the Clade II lineage (Beja-Pereira et al., 2004; Chen et al., 2006; Kimura et al., 2011). Based on archeological evidence, the "Egyptian hypothesis" proposes that villagers in the Nile Valley domesticated the resident Nubian wild ass approximately 5 000–6 000 years ago (Clutton-Brock, 1992; Epstein, 1971); in contrast, the "pastoralist hypothesis", supported by ethnographic, climatic, and linguistic data, states that pastoralists in northeastern Africa 6 500–7 000 years ago domesticated donkeys in response to the increasing aridity of the Sahara (Marshall, 2007). Additionally, the very high genetic diversity found in the Arabian Peninsula, as based on nuclear microsatellites, indicates an alternative domestication center or melting pot (ancient trade areas or routes) (Rosenbom et al., 2015). Ancient DNA also illustrates the extensive distribution of the Nubian and Somali wild ass in Northeast Africa (Kimura et al., 2011), alluding to potential geographic overlaps between the Nubian wild ass and ancestors of the Clade II donkeys. As such, whether these two distinct lineages were domesticated simultaneously or not remains controversial, even though they probably originated from two mitochondrially distinct wild asses (Jordana et al., 2016; Xia et al., 2019).

If the two lineages were domesticated simultaneously, it is reasonable to assume the existence of similar genetic dynamics corresponding to their rapid expansion with innate mobile characteristics. Therefore, in the present paper, we investigated the genetic characteristics of 84 mtDNA genomes and 707 D-loop sequences to compare the dynamic signatures between these two clades that correspond to migration history. Our results will help to further understand mtDNA diversity in African donkeys and also provide novel insights into their domestication histories.

MATERIALS AND METHODS

Sample information

Blood samples were collected and sequenced from animals living in villages or donkey breeding farms all over the world.

This study was approved by the Ethics and Experimental Animal Committee of the Kunming Institute of Zoology, Chinese Academy of Sciences, China. We selected samples without any discrimination toward age, sex, or location. More detailed information, including geographical location and the US National Center for Biotechnology Information (NCBI) GenBank accession Nos., can be found in Supplementary Table S1. Specifically, 5 mL of blood was drawn from healthy animals by a veterinarian (sampling picture in Supplementary Figure S1) with approval from the above-mentioned Ethics and Experimental Animal Committee. The blood samples were preserved in 95% alcohol for further genomic analysis. Whole DNA was isolated from blood using the standard phenol-chloroform method (Ma, 2018) and subsequently quantified with NanoDrop 2000 (Thermo Fisher Scientific, USA).

We obtained a 371 bp length polymerase chain reaction (PCR) product within the D-loop region for 94 donkeys (Tadzhikistan: n=19; Kyrgyzstan: n=5; Kenya: n=16; Nigeria: n=22; Iran: n=30; China: n=2) using primers designed in previous research (Kefena et al., 2014) (5'-DONK-F: CCC AAGGACTATCAAGGAAG-3';DONK-R:5'-GGAATGGCCCTGA AGAAAG-3'). We strictly followed the reaction system as described in Kefena et al. (2014). Specifically, PCR was carried out in a 10 μL reaction volume containing 1 μL of DNA (50 μg/μL), 1 μL 10×buffer (10×Takara Taq^{TM} Buffer Mg^{2+} plus), 0.2 mmol/L of each dNTP, 1 μmol/L of each primer, and 0.2 U/tube Taq (Taq DNA Polymerase, Takara) under the following conditions: initial denaturation at 94 °C for 15 min followed by 45 cycles of denaturation each at 94 °C for 1 min, hybridization at 56 °C for 1 min, extension at 72 °C for 1 min, and final extension at 72 °C for 20 min. The PCR products were gel-purified as described in Ma (2018) and then sequenced by single-strand PCR using the ABI PRISM^{TM} Dye Terminator Cycle Sequencing kit following protocols recommended by the manufacturer. Sequencing was implemented using ABI-PRISM 3730 standard conditions (Applied Biosystems, 2009). For each PCR product, sequences were determined in both forward and reverse directions for all nucleotide positions to avoid possible artificial variations.

Similarly, we obtained the mtDNA genomes of 56 Chinese donkeys and two wild asses using ABI-PRISM 3730 with primers designed in this study (see Supplementary Table S2). Other mtDNA genomes of 17 donkeys (Tadzhikistan: n=6; Kyrgyzstan: n=5; Kenya: n=4; Nigeria: n=2) and two Tadzhikistan horses were yielded through Ion Torrent technology, as per previous study (Chen et al., 2016). Briefly, we amplified PCR fragments, constructed libraries, performed sequencing using next-generation technology, and *de novo* assembled the reads. For quality control, we followed the caveats mentioned in previous mtDNA genome study of domestic animals (Shi et al., 2014). Variants that differentiated from the GenBank reference sequence under accession No. NC_001788 (Xu et al., 1996) were scored. We then manually checked the bam file exported by Torrent Suite 5.0.2 to confirm the scored variants using Integrative Genomics

Viewer (Thorvaldsdóttir et al., 2013).

Collectively, we obtained 94 sequences with a 371 bp length restricted to the D-loop region and 77 almost complete mitochondrial genomes. Sequences generated in this study were deposited in the NCBI archive (GenBank accession Nos.: MK650231–MK650286 for 58 mtDNA genomes obtained by 3730 sequencer; MK619357–MK619411 for 55 D-loop haplotypes of *Equus* yield in this study; SAMN11432793–SAMN11432809 and SRX5702272–SRX5702273 for another 19 mtDNA genomes sequenced through Ion Torrent technology. For specified topology and wider coverage in geographic structure analysis, we combined an additional seven mitochondrial genomes and 536 D-loop sequences downloaded from NCBI website (http://www.ncbi.nlm.nih.gov/), with GenBank accession Nos. listed in Supplementary Table S1. Detailed feature information of all 707 sequences are included in Supplementary Table S1.

Phylogenetic construction and estimation of genetic parameters
We reconstructed the phylogenetic tree of equid mtDNA sequences using neighbor-joining and maximum likelihood analyses with MEGA 6 software (Tamura et al., 2013). For mtDNA genomes, 1 000 bootstrap replicates were conducted with the Kimura 2-parameter model, assuming Gamma distributed mutation rates for neighbor-joining (NJ) analysis and Gamma distributed with invariant sites (G+I) distributed mutation rates for maximum likelihood (ML) analysis. Additionally, we constructed a Bayesian tree using Bayesian evolutionary analysis by sampling trees (BEAST) v1.6.1 (Drummond & Rambaut, 2007) for phylogenetic confirmation with the Gamma distributed HKY substitution model under a strict clock rate, assuming constant size coalescence for tree prior. The non-coding region from 16 129 nt to 16 360 nt was excluded due to the short tandem repeats in both horse and donkey (repeat unit: GTGCACCT in horse; CACACCCACAC ACCCATGCGCGCA in donkey).

We truncated the 77 mitochondrial genomes into the 371 bp length of the D-loop region to ensure combined analysis with the 94 D-loop sequences, which yielded 55 haplotypes. The downloaded 536 D-loop sequences were from different research, which covered different fragments within the D-loop region. Therefore, it was inevitable that a very short overlapping fragment (235 bp) would be yielded as we retrieved as many samples as possible. For genetic diversity analysis, because the downloaded 536 D-loop sequences were haplotype data, we selected only haplotypes within each area from our sequenced samples for parallel comparison. A reduced median-joining network was generated using NETWORK (Bandelt et al., 1999). Average pairwise differences were estimated using Arlequin v3.5 (Excoffier & Lischer, 2010). The nucleotide diversity (π), mismatch distribution, Tajima's D test, and Fu's Fs test were calculated using DnaSP v6 (Rozas et al., 2017).

We assessed the population dynamics for the mtDNA genomes of the two lineages using Bayesian Skyline Plot

(BSP) implemented in BEAST (Drummond & Rambaut, 2007). We estimated the evolution rate using the previously estimated divergence time of 2 million years between the horse and donkey from paleontological data (Lindsay et al., 1980). Migration time to regions distant from the domestication center was inferred by estimating the evolution time of haplotypes only found in the region to the major haplotype.

A geographic distribution map was plotted using the "rworldmap" package (South, 2011) implemented in R (R-Core-Team, 2019).

RESULTS AND DISCUSSION

Two clearly separated domesticated donkey clades

A clearly defined phylogenetic relationship can allow comparison of genetic characters between two lineages. Thus, we constructed a highly supported phylogenetic tree based on the 77 mtDNA genome sequences from the current study (including 73 donkeys from China, Tadzhikistan, Kyrgyzstan, Kenya, and Nigeria, two Asiatic wild asses, and two Tadzhikistan indigenous horses) combined with seven mtDNA genomes downloaded from NCBI (including two donkeys, two Somali wild asses, and two Asian wild asses) (Supplementary Table S1). Using the horse as an outgroup, the Somali wild ass formed a sister clade with all donkeys, a topology highly differentiated from that based on D-loop sequences, which was highly supported by the neighbor-joining, maximum likelihood, and Bayesian methods analyzed here (Figure 1). This topology was also consistent with previous study that concentrated on the major mitochondrial coding regions and placed the Somali wild ass as a sister clade outside all domestic donkeys with high confidence, although they did not label the lineage information (Sun et al., 2016). Here, we obtained consistent topologies for the major clades according to both whole mitochondrial genomes and coding regions only (Figure 1; Supplementary Figure S2). With a view that coding regions experience restricted selective sweep, whereas selective force is potentially relaxed on the D-loop regions (Endicott et al., 2009), it is reasonable to assume that the phylogenetic relationship revealed by mitochondrial genome data would be much closer to reality.

The two donkey lineages were clearly separated from each other, with high support (Figure 1), further confirming the existence of two lineages. Considering the lack of geographic representation from mtDNA genomes (58 of 75 domestic donkeys were from China), we reconstructed the phylogenetic tree using both neighbor-joining and maximum likelihood methods with D-loop sequences from the 171 sequences obtained in this study and 536 sequences publicly available on NCBI (see sample information in Supplementary Table S1 and phylogenetic tree in Supplementary Figure S3). Similarly, the donkeys were divided into two groups. Although the clades showed lower support, which is common in analysis of short segments like the D-loop region, all sequences with published lineage information clustered into their corresponding clades. In addition, the 73 donkey D-loop sequences obtained in this study showed consistent lineage sorting, as observed in the whole mtDNA genome phylogeny, further supporting the D-loop region phylogenetic relationship. Our phylogenetic tree based on D-loop sequences also revealed a sister relationship between the Somali wild ass and all domesticated donkeys (Supplementary Figure S3), consistent with the topology revealed by the mitochondrial genome, but inconsistent with previous studies showing different topology using D-loop sequences (Beja-Pereira et al., 2004; Han et al., 2014). The number of haplotypes used for phylogenetic construction may account for this discrepancy, as more haplotypes represent greater genetic diversity, and subsequently more statistical credibility. To test this, we randomly selected donkey haplotypes of equal sample size, as previously described in Han et al. (2014), and constructed a phylogenetic tree with the same parameters in this study. We obtained similar phylogenetic topology (Supplementary Figure S4), showing a closer relationship shared by the Somali wild ass and domesticated Clade II lineage. Therefore, a large sample size for phylogenetic analysis can compensate for the limited information obtained from short segments.

Different demographic dynamics of two lineages

As the donkey samples were credibly sorted, we subsequently compared several population characteristics between the two lineages. First, we scanned sequence variation within the mitochondrial genome for each lineage. The Clade I lineage showed a total of 288 varied sites: 258 in the coding region and 30 in the non-coding regions (D-loop). The Clade II lineage showed a total of 188 varied sites: 168 in the coding region and 20 in the non-coding regions (D-loop). Figure 2A illustrates the distribution of nucleotide diversity (π) within each lineage along the genome based on assessment of 200 nt windows (step size=100 nt) centered at the midpoint. Although the highest diversity was observed in the D-loop region presenting the short tandem repeats of CACACCCACACACCCATGCGCGCA (donkey reference NC_001788: from 16 173 nt to 16 340 nt) in both lineages, the Clade I lineage possessed significantly more segregation sites (average segregation sites per window=1.300) in the coding region than the Clade II lineage (average segregation sites per window=1.091) (Wilcox test: P=0.016), whereas no significant differences were found in segregation counts in the non-coding region in the two lineages (average segregation sites per window in Clade I and Clade II were 3.057 and 2.303, respectively; Wilcox test: P=0.628). This discrepancy may be due to the saturation effect in the D-loop region, demographics, selective forces, and/or other factors suggestive of potential independent domestic histories between the two lineages. To further address this issue, we used mtDNA genomes to reconstruct ancestral population dynamics with BSP for each lineage (Figure 2B). A similar demographic history would be expected under the assumption of simultaneous domestication. However, we observed a constant effective population size in the Clade II lineage during most history, whereas the Clade I lineage experienced an apparent population increase approximately 8 000 years
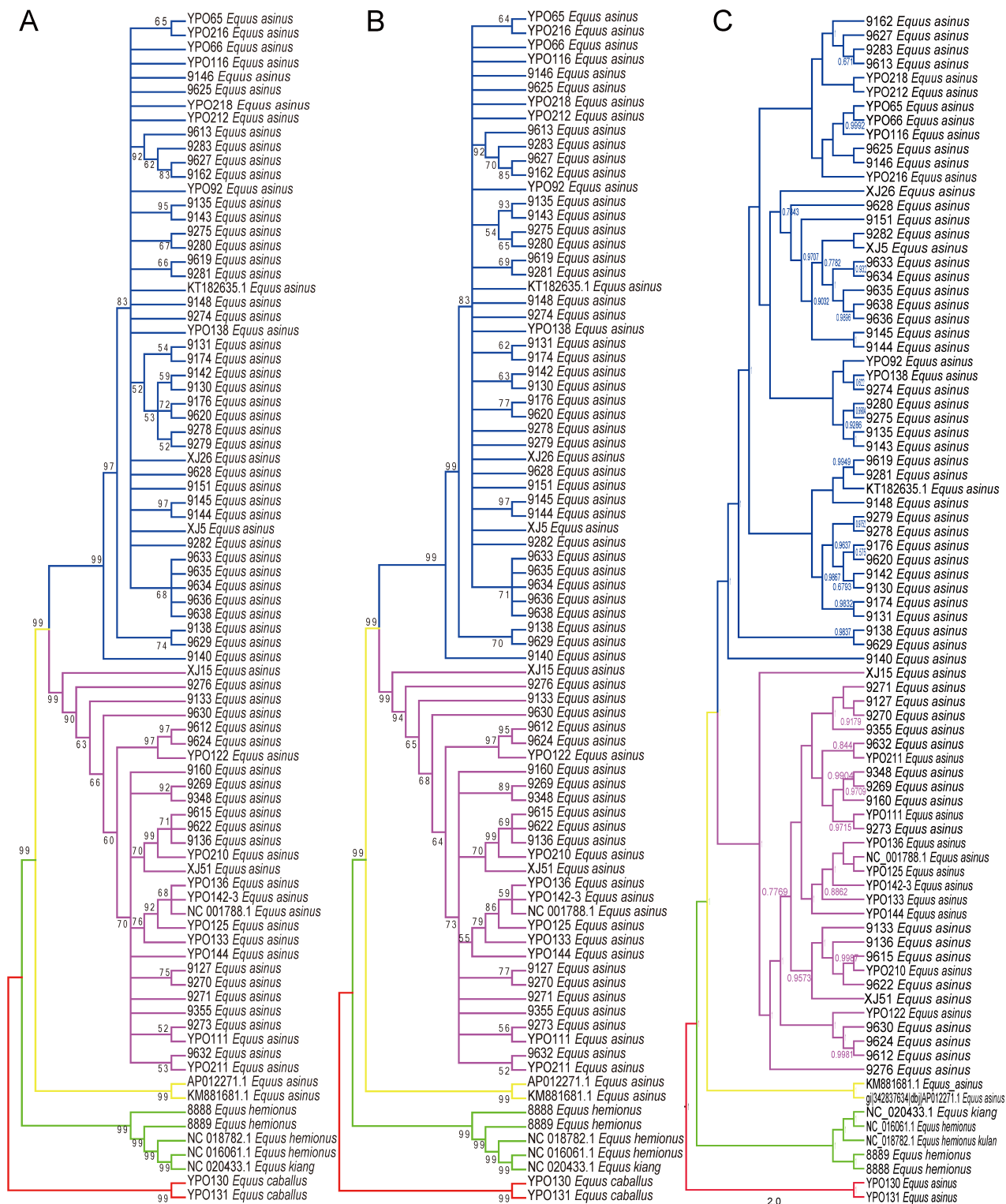
Figure 1 mtDNA genome phylogeny of *Equus*

A: Neighbor-joining tree; B: Maximum likelihood tree; C: Bayesian tree. Colored branches denote different categories. Red: Horse; Green: Asiatic wild ass; Yellow: Somali wild ass; Blue: Clade I lineage; Magenta: Clade II lineage.

ago, coinciding with the archeological date 7 770±95 before present (BP) (Marshall, 2007). Noticeably, both lineages exhibited a marked decrease in recent years, probably due to the occurrence of the industrial revolution. To further confirm this discrepancy in population dynamics, we assessed the demographic history of the two lineages based on D-loop sequences, with nucleotide mismatch distribution (Rogers & Harpending, 1992), Fu's Fs test, and Tajima's D test. The nucleotide mismatch curve showed a single peak in the Clade

I lineage and double peaks in the Clade II lineage (Figure 2C). Moreover, results showed that both Fu's Fs test and Tajima's D test significantly deviated from neutrality in the Clade I lineage (Tajima's D=−2.310, P<0.01; Fu's Fs=−33.909) but not in the Clade II lineage (Tajima's D = −0.993, P >0.10; Fu's Fs= −17.464). These results were in accordance with the expansion in the Clade I lineage and constant size in the Clade II lineage, confirming the demographic dynamics revealed by mitochondrial genomes.

Figure 2 Population characteristics between Clade I and Clade II lineages

A: Sequence variation of mtDNA genome within each lineage. Nucleotide diversity (π) using 200 nt windows (step size=100 nt) centered at midpoint. Schematic (linearized) genetic map of mitochondrial genome is presented. B: Demographic histories for each lineage estimated using Bayesian skyline plot from mitochondrial genomes. Shadow means 95% highest posterior density intervals of effective number of females. C: Mismatch distribution for each mtDNA lineage. D: Median-joining network based on D-loop region. Red diamonds are haplotypes of horse outgroup. Green circles are haplotypes of Asiatic wild ass. Dark red circles are haplotypes of Somali wild ass. Indigo circles are haplotypes of Clade I lineage. Magenta circles are haplotypes of Clade II lineage. Circle size is proportional to frequencies of haplotypes.

Discrepancies of haplotype structure

In addition to demographic estimation, we also investigated several other genetic characters to assess whether they showed similar evolution patterns between the two lineages. We first constructed a reduced median network based on D-loop sequences (Figure 2D). Similar to previous study (Beja-

Pereira et al., 2004; Kimura et al., 2011), the Clade II haplotypes were mainly derived from a single major haplotype (H16), with a simple star-like shape, whereas the genetic architecture of the Clade I haplotypes was more complicated, with more universally occurring haplotypes (e.g., H20, H52, H22). This was consistent with the much higher genetic distance and nucleotide diversity within the Clade I lineage (average pairwise distance=1.932 7, SD=1.100 7; Pi=0.008 4, SD=0.000 3) than within the Clade II lineage (average pairwise distance=1.074 4, SD=0.711 2; Pi=0.004 6, SD=0.000 5), implying that the Clade I lineage involved many more individuals at the beginning of domestication compared with the Clade II lineage (Table 1).

Table 1 Genetic distance and nucleotide diversity of two clades

	Pairwise distance	SD (distance)	Pi	SD (Pi)
Clade I lineage	1.932 7	1.100 7	0.008 4	0.000 3
Clade II lineage	1.074 4	0.711 2	0.004 6	0.000 5

Biased distribution of two lineages in Africa

Independent migration events may result in a geographical structure. Therefore, we estimated the migration routes of the two lineages by referring to their proportion across the world. D-loop sequences sampled from the Balkans and microsatellites sampled from northeast Africa, the Near East, and the Arabian Peninsula indicated the absence of a geographical structure (Pérez-Pardal et al., 2014; Rosenbom et al., 2015). Indeed, samples collected from Middle Asia (Pakistan, Kazakhstan) and major areas of China demonstrated an almost equal proportion of the two lineages (Figure 3A). The Arabian Peninsula is assumed to be the melting pot from where domestic donkeys migrated to the world (Rosenbom et al., 2015). As this area showed a nearly equal proportion of the two lineages, it is reasonable that a similar pattern is commonly observed across the Eurasian mainland (Beja-Pereira et al., 2004). Nevertheless, all 20 samples collected from Nigeria belonged to the Clade I lineage. When we focused on the distribution in Africa, an apparent spatial structure was detected: donkeys from sub-Saharan Africa (e.g., Ghana, Guinea, Benin, Mali, Senegal, South Africa, and Burkina Faso) tended to be descended from the Clade I lineage, whereas the Clade II lineage was dominant along the East (e.g., Eritrea, Somalia, Swaziland, and Zambia) and North coasts (e.g., Libya, Tunisia, and Morocco) (Figure 3A).

The time of migration to areas distant from the domestication center can be inferred by dating the time to the major haplotype for the derived haplotypes only found in that area. In this way, we inferred that the Clade I lineage migrated into sub-Sahara 4 874.28±1 817.92 years ago, around the commencement of desertification in the Sahara (5 000 to 7 000 BP) (Marshall, 2000). Due to the excellent tolerance of donkeys for deserts, it is reasonable to assume that donkeys, rather than horses, were the major means of transport across the Sahara during the initial period of desertification. Given the biased Clade I lineage distribution in the sub-Sahara, this

migration time provides possible evidence for the "pastoralist hypothesis": i. e., pastoralists in northeastern Africa domesticated Clade I lineage donkeys in response to the increasing aridity in the Sahara.

Donkeys are thought to be have been brought into Europe by the second millennium BC, possibly through viticulture introduction, as the donkey is associated with the Syrian god of wine, Dionysus (Meutchieye et al., 2017). Interestingly, the majority of sequences from the Iberian Peninsula belonged to the Clade II lineage, much different from those collected in other parts of Europe (Figure 3A). Additionally, the estimated time of arrival in the Iberian Peninsula was 5 336.8±1 652.56 years ago, much earlier than the known history of ~4 000 years (Cardoso et al., 2013). Considering the dominance of the Clade II lineage in Morocco, it could be assumed that donkeys migrated into the Iberian Peninsula directly through the Strait of Gibraltar. Furthermore, the unequal distribution of Clade I and Clade II in Europe may account for the pronounced footprint in American following the complex process of colonization, where an apparent geographic structure has been observed (Jordana et al., 2016; Xia et al., 2019). The Clade II lineage also migrated dominantly along East Africa, implying a potential association with ancient social expansion, such as the Bantu expansion (Hiernaux, 1968; Herrera & Garcia-Bertrand, 2018), and ancient trade routes of empires such as the Punt and Aksum (Andrews, 2017; Mark, 2011; Mukhtār, 1981).

Different routes of two lineages during expansion to China

As populations expand from centers of origin, genetic diversity is lost as a consequence of the limited numbers of individuals involved in the expansionist movement. Although no apparent geographic structure was detected across China, where the two lineages show an almost equal proportion (Figure 3A), discrepancies in diversity decay may exist between the two lineages if non-simultaneous expansion into China occurred, thereby suggesting possible private migration routes. The Xinjiang Province may be a transportation center as it demonstrated almost the highest nucleotide diversity for both lineages (Figure 3B, C), consistent with previous studies on genetic diversity among various Chinese breeds (Ge et al., 2007) and written records (Xie, 1987). The genetic diversity of Clade I remained relatively high in the Qinghai and Henan provinces, but declined gradually towards north and south, respectively (Figure 3B), consistent with a migration route from Xinjiang to the Guanzhong Plain through the Ningxia and Gansu provinces, and finally other areas of China inferred previously (Ge et al., 2007). The genetic diversity of Clade II remained relatively high in the Inner Mongolia and Yunnan provinces, but declined substantially in the middle region of China (Figure 3C), consistent with previously inferred migration routes from Xinjiang to Inner Mongolia (north forward) and to Yunnan (south forward) (Ge et al., 2007). These potential migration routes partially match written records, which suggest that the "Taihang donkey" was introduced into Hebei Province through Inner Mongolia; the "Yunnan donkey" was introduced from Xinjiang to Yunnan, as
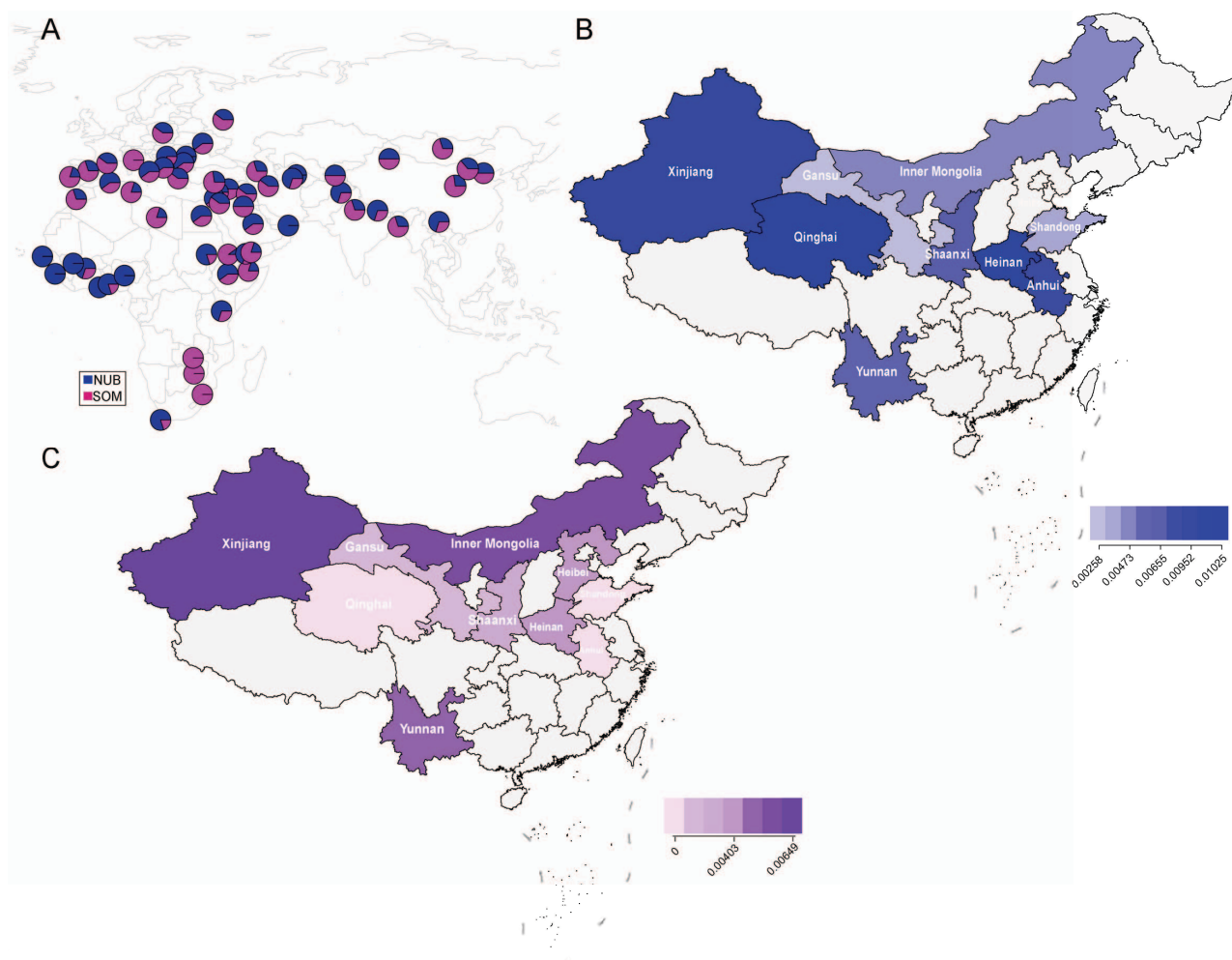
Figure 3 Worldwide geographic distribution of two lineages plotted using "rworldmap" package in R (A) and genetic diversity of Clade I (B) and Clade II (C) lineages across China
A: Blue: Clade I lineage; Magenta: Clade II lineage. B: Dark to light indigo indicates diversity decay. Gray regions indicate no samples available. C: Dark to light magenta indicates diversity decay. Gray regions indicate no samples available.

it shares many morphological traits with the Xinjiang donkey; and many other breeds may have been formed through migration routes from Xinjiang to Gansu and then Guanzhong Plain, Henan, and other areas of China (Xie, 1987). Therefore, it is likely that the two lineages expanded into China through different migration events, although from the same transportation center (Xinjiang).

CONCLUSIONS

Overall, our results on the demographic dynamics, haplotype structure, diversity decay pattern, and distribution bias in Africa are in accordance with non-simultaneous domestication events and an independent migration history. Our findings revealed that domestication of the donkey may have been driven by the response of pastoralists to the desertification of the Sahara and by social expansion and trade of ancient humans. Future research on population genomes will be

needed to increase our understanding of donkey domestication history.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y. P. Z., Y. L., and S. C. O. designed the study. X. Y. M., T. N., and A. C. A. sequenced the samples. X.Y.M. and J.L. performed data analyses. A.E., J. D.L., B.R.A., J.I., A.A.A., S.F.W., H.Q.L., N.T.A., M.E.A., K.B.I., R.A.M.A., S. C.O., O.J.S., M.G.F., X.C., and W.K.Y. helped with sample collection. Y.L., A.C.A., Z.W., and M.S.P. wrote the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Andrews E. 2017. 7 Influential African Empires. From ancient Sudan to medieval Zimbabwe, get the facts on seven African kingdoms that made their mark on history. A&E Television Networks, LLC. URL: https://www.history.com/.amp/news/7-influential-african-empires.

Applied Biosystems. 2009. DNA Sequencing by Capillary Electrophoresis, Applied BiosystemsChemistry Guide. Thermo Fisher Scientific.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**(1): 37−48.

Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhiet AO, Abdalla MA, Mashkour M, Jordana J, Taberlet P, Luikart G. 2004. African origins of the domestic donkey. *Science*, **304**(5678): 1781.

Blench RM. 2000. A History of Donkeys, Wild Asses and Mules in Africa. London, UK: University College London Press.

Cardoso, J L, Vilstrup JT, Eisenmann V, Orlando L. 2013. First evidence of *Equus asinus* L. in the Chalcolithic disputes the Phoenicians as the first to introduce donkeys into the Iberian Peninsula. *Journal of Archaeological Science*, **40**(12): 4483−4490.

Chen SY, Zhou F, Xiao H, Sha T, Wu SF, Zhang YP. 2006. Mitochondrial DNA diversity and population structure of four Chinese donkey breeds. *Animal. Genetics*, **37**(4): 427−429.

Chen X, Ni G, He K, Ding ZL, Li GM, Adeola AC, Murphy RW, Wang WZ, Zhang YP. 2016. An improved de novo pipeline for enrichment of high diversity mitochondrial genomes from Amphibia to high-throughput sequencing. *BioRxiv*: 080689.

Clutton-Brock J. 1992. Horse Power: A History of the Horse and the Donkey in Human Societies. Cambridge, USA: Harvard University Press.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary. Biology*, **7**: 214.

Endicott P, Ho SYW, Metspalu M, Stringer C. 2009. Evaluating the mitochondrial timescale of human evolution. *Trends in Ecology and Evolution*, **24**(9): 515−521.

Epstein H. 1971. The Origin of the Domestic Animals of Africa. New York, USA: African Publishing Corporation (APC).

Excoffier L, Lischer HEL. 2010. Arlequin suite ver3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**(3): 564−567.

Ge QL, Lei CZ, Jiang YQ, Chen H, Zhang W, Dang RH, Zhen HL, Zhang AL, Li TP. 2007. Genetic diversity and origins of Chinese donkey revealed by mitochondrial D-loop sequences. *Acta Veterinaria et Zootechnica Sinica*, **38**(7): 641−645.

Han L, Zhu S, Ning C, Cai D, Wang K, Chen Q, Hu S, Yang J, Shao J, Zhu H, Zhou H. 2014. Ancient DNA provides new insight into the maternal lineages and domestication of Chinese donkeys. *BMC Evolutionary Biology*, **14**: 246.

Herrera RJ, Garcia-Bertrand R. 2018. The Bantu Expansion. Ancestral DNA, human origins, and migrations. Academic Press, 395–432.

Hiernaux J. 1968. Bantu expansion: the evidence from physical anthropology confronted with linguistic and archaeological evidence. *The Journal of African History*, **9**(4): 505−515.

Jordana J, Ferrando A, Miró J, Goyache F, Loarca A, Martinez López OR, Canelón JL, Stemmer A, Aguirre L, Lara MAC, Álvarez LA, Llambí S, Gómez N, Gama LT, Nóvoa MF, Martínez RD, Pérez E, Sierra A, Contreras MA, Guastella AM, Marletta D, Arsenos G, Curik I, Landi V, Martinez A, Delgado JV. 2016. Genetic relationships among American donkey populations: insights into the process of colonization. *Journal of Animal Breeding and Genetics*, **133**(2): 155−164.

Kefena E, Dessie T, Tegegne A, Beja-Pereira A, Yusuf Kurtu M, Rosenbom S, Han JL. 2014. Genetic diversity and matrilineal genetic signature of native Ethiopian donkeys (*Equus asinus*) inferred from mitochondrial DNA sequence polymorphism. *Livestock Science*, **167**: 73−79.

Kimura B, Marshall FB, Chen S, Rosenbom S, Moehlman PD, Tuross N, Sabin RC, Peters J, Barich B, Yohannes H, Kebede F, Teclai R, Beja-Pereira A, Mulligan CJ. 2011. Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. *Proceedings of the Royal Society B: Biological Sciences*, **278**(1702): 50−57.

Lindsay EH, Opdyke ND, Johnson NM. 1980. Pliocene dispersal of the horse Equus and late Cenozoic mammalian dispersal events. *Nature*, **287**: 135−138.

Ma XY. 2018. Global Perspective of Phylogeography and Demongraphic History of *Equus asinus* Based on Mitochondrial Analysis. Master thesis, Yunnan University, China.

Mark JJ. 2011. Punt. Ancient History Encyclopedia. https://www.ancient.eu/punt/.

Marshall F. 2000. The Origins and Development of African Livestock. London: University College London Press.

Marshall F. 2007. Rethinking Agriculture: Archeological and Ethnoarcheological Perspectives. Walnut Creek, CA, USA: Left Coast Press.

Meutchieye F, Kwalar NN, Nyock RAF. 2017. Donkey Husbandry and Production Systems. Current Donkey Production and Functionality · Relationship with Humans. F. J. N. González, UCO Press.

Mukhtār MJa-D . 1981. Ancient Civilizations of Africa. London: Heinemann Educational Books.

Pérez-Pardal L, Grizelj J, Traoré A, Cubric-Curik V, Arsenos G, Dovenski T, Marković B, Fernández I, Cuervo M, Álvarez I, Beja-Pereira A, Curik I, Goyache F. 2014. Lack of mitochondrial DNA structure in Balkan donkey is consistent with a quick spread of the species after domestication. *Animal Genetics*, **45**(1): 144−147.

R-Core-Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**(3): 552−569.

Rosenbom S, Costa V, Al-Araimi N, Kefena E, Abdel-Moneim AS, Abdalla MA, Bakhiet A, Beja-Pereira A. 2015. Genetic diversity of donkey populations from the putative centers of domestication. *Animal Genetics*, **46**(1): 30−36.

Rossel S, Marshall F, Peters J, Pilgram T, Adams MD, O'Connor D. 2008. Domestication of the donkey: timing, processes, and indicators. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(10): 3715−3720.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, **34**(12): 3299−3302.

Shi NN, Fan L, Yao YG, Peng MS, Zhang YP. 2014. Mitochondrial genomes of domestic animals need scrutiny. *Molecular Ecology*, **23**(22): 5393−5397.

Smith DG, Pearson RA. 2005. A review of the factors affecting the survival of donkeys in semi-arid regions of Sub-Saharan Africa. *Tropical Animal Health and Production*, **37**(S1): 1−19.

South A. 2011. Rworldmap: a new R package for mapping global data. *The R Journal*, **3**(1): 35−43.

Starkey P. 2000. The History of Working Animals in Africa. London, UK: University College London Press.

Sun Y, Jiang Q, Yang C, Wang X, Tian F, Wang Y, Ma Y, Ju Z, Huang J, Zho X u, Zhong J, Wang C. 2016. Characterization of complete mitochondrial genome of Dezhou donkey (*Equus asinus*) and evolutionary analysis. *Current Genetics*, **62**(2): 383−390.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version6.0. *Molecular Biology and Evology*, **30**(12): 2725−2729.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings Bioinformatics*, **14**(2): 178−192.

Xia X, Yu J, Zhao X, Yao Y, Zeng L, Ahmed Z, Shen S, Dang R, Lei C. 2019. Genetic diversity and maternal origin of Northeast African and South American donkey populations. *Animal Genetics*, **50**(3): 266−270.

Xie CX. 1987. The *Equus* Breeds in China. Shanghai: Shanghai Science & Technical Publishers. (in Chinese)

Xu X, Gullberg A, Arnason U. 1996. The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *Journal of Molecular Evolution*, **43**(5): 438−446.

First genetic assessment of brackish water polychaete *Tylorrhynchus heterochaetus*: mitochondrial *COI* sequences reveal strong genetic differentiation and population expansion in samples collected from southeast China and north Vietnam

Xing-Han Chen[1,#], Sen Yang[2,#], Wei Yang[1], Yuan-Yuan Si[1], Rui-Wen Xu[1], Bin Fan[1], Le Wang[3], Zi-Ning Meng[4,5,*]

[1] *Department of Food and Environmental Engineering, Key Laboratory for Marine Estuary Fishery Resources Protection of Yangjiang City, Yangjiang Polytechnic, Yangjiang, Guangdong 529500, China*

[2] *College of Animal Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, Guangdong 520175, China*

[3] *Temasek Life Sciences Laboratory, National University of Singapore, Singapore 117604, Singapore*

[4] *State Key Laboratory of Biocontrol, Institute of Aquatic Economic Animals, and the Guangdong Province Key Laboratory for Aquatic Economic Animals, Life Science School, Sun Yet-Sen University, Guangzhou, Guangdong 510275, China*

[5] *Southern Laboratory of Ocean Science and Engineering, Zhuhai, Guangdong 519000, China*

ABSTRACT

Tylorrhynchus heterochaetus is a widespread benthic polychaete worm found in coastal brackish waters of the west Pacific. It has high ecological and economic value as a biomarker of water quality and as a high-quality feed in aquaculture and fisheries and is considered a delicacy in some areas of Asia. However, it has experienced a marked reduction in recent years due to overexploitation as well as changes in the environment and climate. Here, to comprehensively understand its genetic background and thus provide insights for better conservation and utilization of this species, we assessed the genetic variability and demographic history of *T. heterochaetus* individuals sampled from eight locations along the coasts of southeast China and

north Vietnam based on mitochondrial cytochrome c oxidase I (*COI*) sequences. We observed high haplotype diversity (*Hd*), with an average of 0.926, but relatively low nucleotide diversity (π), with a mean of 0.032 across all samples. A total of 94 polymorphic sites and 85 haplotypes were identified among 320 individuals. The pairwise genetic distances among haplotypes ranged from 0.001 to 0.067, with the high intraspecific divergence possibly reflecting geographic isolation and gene pool fragmentation. Significant genetic structures were revealed among the studied locations; specifically, the eight locations could be treated as six genetically

different populations based on pairwise Φ_{ST} results (0.026–0.951, P <0.01). A significant pattern of isolation-by-distance was detected between the genetic and geographic distances (r=0.873, P=0.001). Three geographic lineages were defined based on phylogenetic tree and network analyses of *COI* haplotypes. AMOVA results indicated that genetic variations mainly occurred among the three lineages (89.96%). Tests of neutrality and mismatch distribution suggested that *T. heterochaetus* underwent recent population expansion. These results provide the first report on the genetic status of *T. heterochaetus* and will be valuable for the management of genetic resources and better understanding of the ecology and evolution in this species.

INTRODUCTION

Tylorrhynchus heterochaetus, a member of the polychaete family Nereididae (Annelida: Phyllodocida), is a widespread benthic invertebrate found in brackish waters along the coasts of China, Japan, and Southeast Asia (Tuan, 2018). Due to its high sensitivity to water quality, the species is widely used as biomarker of marine environmental conditions (Dean, 2008). Moreover, it has great potential in both aquaculture and recreational fisheries as a high-quality feed (Costa et al., 2006) and is also a favored and relatively expensive delicacy in some areas of Asia, such as Vietnam and southern China (Glasby & Timm, 2008).

During the breeding season, *T. heterochaetus* adults tend to swim to higher salinity waters and aggregate for spawning, and then die after releasing gametes. The salinity levels suitable for reproduction range from 10 to 13 ppt, and the hatching rate can be significantly affected by different salinities (Duan et al., 2017). For further development, however, early setiger larvae (3–5 d after hatching) prefer a low salinity environment, where they settle into mud until they reach sexual maturity in the following year. Thus, *T. heterochaetus* individuals are usually confined to the muddy bottom in brackish water estuaries and are thus extremely vulnerable to the impact of sudden changes in the external environment. In recent years, the increasing demand for commercial utilization, together with changes in both climate and environment, such as water pollution, ocean salinization, and habitat fragmentation, has led to a marked reduction in the natural resources of *T. heterochaetus*, even in historically high-yield habitats (Tuan, 2018).

Based on the above threats, a comprehensive understanding of the genetic background of *T. heterochaetus* could greatly facilitate its conservation and management and thus better utilization of its genetic resources. To date,

however, studies on *T. heterochaetus* have mainly focused on reproductive physiology (Kaoawa, 1954; Okada, 1952; Sato & Osanai, 1990; Tuan, 2018) and the function of individual genes (Green et al., 1995; Suzuki & Gotoh, 1986; Suzuki et al., 1990). As such, the genetic diversity and structure of natural *T. heterochaetus* populations remain unclear.

Mitochondrial DNA (mtDNA) is an effective tool for molecular phylogenesis and population genetics analysis of polychaetes, such as *Marenzelleria* (Blank & Bastrop, 2009), *Phascolosoma esculenta* (Gao et al., 2018), *Perinereis aibuhitensis* (Liu et al., 2012), *Pygospio elegans* (Kesäniemi et al., 2012), *Aglaophamus australiensis*, and *Nephtys longipes* (Smith et al., 2015), due to the advantages of maternal inheritance, relatively fast mutation rate, and non-recombination (Birky et al., 1989). In the present study, we used the cytochrome *c* oxidase subunit I (*COI*) gene from mtDNA to investigate the genetic diversity, population structure, and demographic history of *T. heterochaetus* along the coasts of southeast China and north Vietnam. This genetic survey will provide useful information for the development of effective conservation and utilization strategies of this species.

MATERIALS AND METHODS

Sample collection and DNA extraction

A total of 320 individuals from eight locations were collected during 2016–2018 from the coastal waters of southeast China and north Vietnam. Two locations were along the coast of Vietnam (Hai Phong (HP), Nam Định (ND)) and six locations were along the coast of China (Wenzhou (WZ), Fu'an (FA), Fuzhou (FZ), Yangjiang (YJ), Zhongshan (ZS), Qinzhou (QZ)). Detailed geographic locations and sampling information are shown in Figure 1 and Table 1. The samples were stored in 95% ethyl alcohol at −20 °C until DNA extraction. Muscle tissues (~10 mg for each individual) were dissected and genomic DNA was extracted using a Genomic DNA Extraction Kit (Tiangen Biotech, DP304, Beijing, China) according to the manufacturer's protocols. The extracted DNA was stored at −80 °C before use.

Gene amplification and sequencing

Partial sequences of mtDNA *COI* were amplified and sequenced using universal DNA primers *COI*-LCO1490 and HCO2198 (Folmer et al., 1994). Polymerase chain reaction (PCR) amplification was conducted in a 50 μL volume containing 0.5 μmol/L of each primer, 0.2 mmol/L of each dNTP, 30 ng of template DNA, 1.5 mmol/L MgCl2, 1×PCR buffer, and 1 unit of Taq DNA polymerase (Fermentas, Thermo Scientific, USA). PCR was conducted with an initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 45 s, with a final extension at 72 °C for 3 min. The amplified products were then purified and sequenced on an ABI Prism 3730 DNA sequencer (Applied Biosystems, USA) using both forward and reverse primers individually.

Data analysis

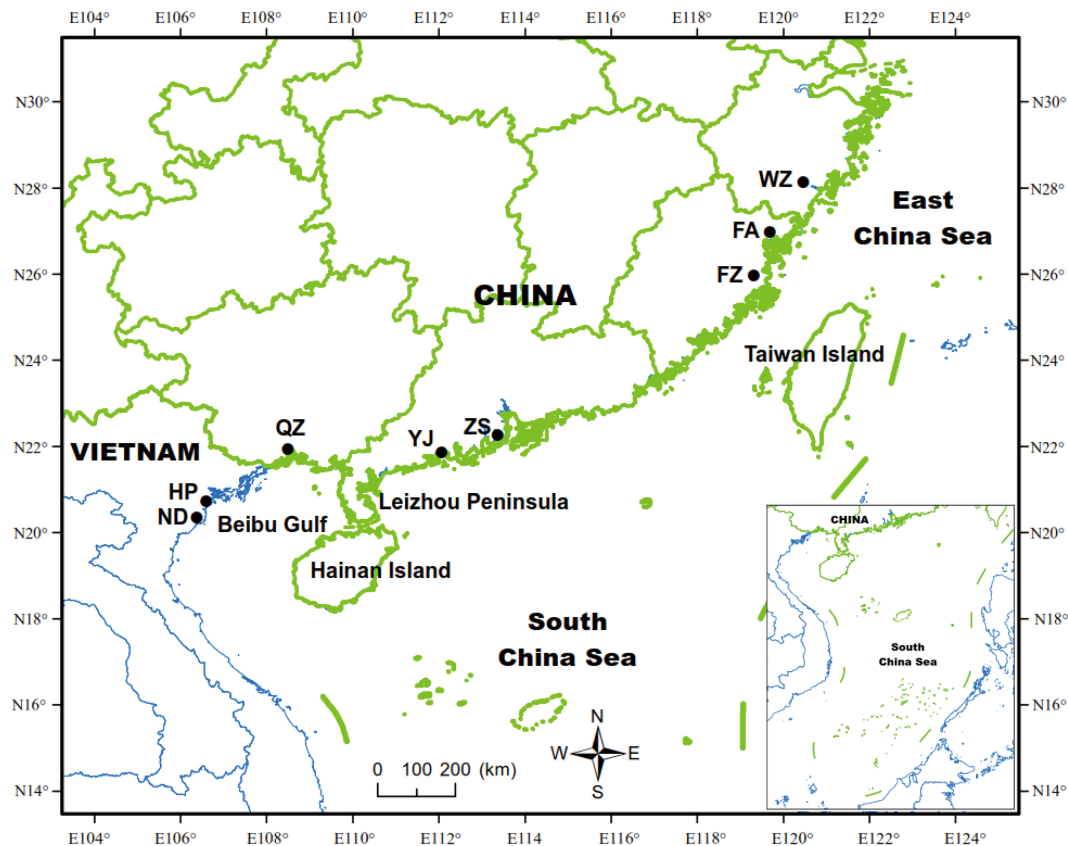Using the *COI* gene from the complete mitochondrial genome

Figure 1 Sampling locations of *T. heterochaetus* used in this study (indicated by solid cycles)

Corresponding location abbreviations are as in Table 1.

Table 1 Sampling information on *T. heterochaetus* in this study

Location code	Locality	Country	Sample size (*n*)	Longitude (E)	Latitude (N)
WZ	Wenzhou, Zhejiang	China	40	121°18'	28°38'
FA	Fu'an, Fujian	China	40	119°40'	26°59'
FZ	Fuzhou, Fujian	China	40	119°18'	25°58'
YJ	Yangjiang, Guangdong	China	40	112°02'	21°51'
ZS	Zhongshan, Guangdong	China	40	113°21'	22°15'
QZ	Qinzhou, Guangxi	China	40	108°29'	21°56'
HP	Hai Phong	Vietnam	40	106°35'	20°43'
ND	Nam Định	Vietnam	40	106°22'	20°20'

sequence of *T. heterochaetus* as a reference sequence (Chen et al., 2016), sequence data were aligned using Cluster X 2.0 (Thompson et al., 1997). DNA sequence polymorphisms, including number of polymorphic sites (*S*), number of haplotypes (*H*), haplotype diversity (*Hd*), nucleotide diversity (*π*), and average number of nucleotide differences (*K*), were estimated using DnaSP 5.10 (Librado & Rozas, 2009). To examine the genealogical relationships among mtDNA haplotypes, a haplotype network was constructed based on the median-joining algorithm in Network 5.0 (Bandelt et al., 1999). A neighbor-joining phylogenetic tree of *COI* haplotypes was constructed in MEGA 7.0 (Kumar et al., 2016) with the Kimura-2-parameter model (Saitou & Nei, 1987) and 1 000

bootstrap replicates. *Perinereis aibuhitensis*, a sea worm from the family Nereididae, was used as the outgroup (GenBank accession No.: NC023943.1). Pairwise genetic distances among haplotypes or different locations of *T. heterochaetus* were also calculated using MEGA 7.0 based on the Kimura-2-parameter model.

Pairwise Φ_{ST}, analysis of molecular variance (AMOVA), as well as correlation between genetic and geographic distance (coastline distance between sampling sites measured by Google Earth) estimated with the Mantel test, were all calculated in Arlequin v3.5 with 10 000 permutations (Excoffier et al., 2005). For examining demographic history of *T. heterochaetus*, Tajima's *D* (Tajima, 1989) and Fu's *F*s (Fu,

1997) tests were used to examine the neutrality of coding sequences with Arlequin v3.5. Additionally, mismatch distribution analysis was also performed to investigate the hypothesis of population expansion. The fitness between the observed and simulated distributions was tested using the sum of squared deviations (*SSD*) as well as Harpending's raggedness index (*HRI*) (Harpending, 1994). Time of population expansion was calculated with the formulas Tau (τ)=2ut and u=2μk, where t is the time since expansion, u is the mutation rate for the *COI* gene (Rogers & Harpending, 1992), k is the number of nucleotides, and μ is the mutation rate. In this study, a mutation rate of 2% per million years was used as suggested by Olson et al. (2009) for another polychaete *Hobsonia florida*.

RESULTS

Genetic diversity

A 709 bp fragment of the *COI* gene was obtained after alignment. The average nucleotide frequency was: T=29.7%, C=24.4%, A=29.1%, and G=16.9%. Of the 320 sequences from the eight locations, a total of 94 variable sites were observed, including 32 singleton variable sites and 62 parsimony informative sites. Genetic diversity indices are presented in Table 2. A total of 85 haplotypes (GenBank accession No.: MK614603−MK614686) were identified, most of which (68 out of 85) were represented by a single individual, and pairwise genetic distances among haplotypes varied from 0.001 to 0.067. The number of haplotypes at a location ranged from 8 to 23. Overall, most locations showed moderate to high haplotype diversity (0.237−0.949) due to the large number of rare haplotypes. However, nucleotide diversity was relatively low, ranging from 0.000 4 to 0.009 12, across all studied locations. The WZ location exhibited the highest haplotype diversity (*Hd*=0.949) and nucleotide diversity (π=0.009).

Population structure and phylogenetic analysis

Pairwise Φ_{ST} values ranged from −0.013 to 0.951 for all locations and were highly significant (*P*<0.01), except for two

Table 2 Genetic diversity of eight *T. heterochaetus* locations based on *COI* gene sequences

Location (lineage)	*n*	*S*	*H*	*Hd*	π	*k*
WZ	40	32	23	0.949	0.009 12	6.465
FA	40	13	10	0.776	0.004 74	3.362
FZ	40	16	13	0.783	0.004 70	3.329
ZS	40	18	6	0.237	0.001 27	0.900
YJ	40	12	13	0.654	0.001 40	0.994
QZ	40	11	10	0.441	0.001 04	0.735
HP	40	17	16	0.645	0.001 59	1.124
ND	40	7	8	0.618	0.001 05	0.744
Lineage A	120	42	39	0.861	0.007 11	5.043
Lineage B	80	29	18	0.660	0.001 75	1.239
Lineage C	120	29	29	0.767	0.002 74	1.942
Overall	320	94	85	0.926	0.032 15	22.794

n: Sample size; *H*: Number of haplotypes; *S*: Number of polymorphic sites; *Hd*: Haplotype diversity; π: Nucleotide diversity; *k*: Mean number of pairwise differences. WZ: Wenzhou; FA: Fu'an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Định. Lineage A=WZ+FA+FZ, Lineage B=ZS+YJ, Lineage C=QZ+HP+ND.

close location pairs (FA-FZ and HP-ND), whereas pairwise genetic distance varied from 0.001 to 0.063 (Table 4). Thus, all eight locations sampled could be treated as six genetically different populations (WZ, FA+FZ, ZS, YJ, QZ, HP+ND, Table 4). The topologies produced from both the haplotype phylogenetic tree (Figure 2) and haplotype median-joining network (Figure 3) showed a consistent structure. All eight locations could be characterized into three geographically distinguishable lineages (i.e., lineage A: WZ, FA+FZ; lineage B: ZS, YJ; lineage C: QZ, HP+ND, Figures 1−3). The AMOVA results based on the three lineages revealed that genetic variation mainly occurred among lineages (89.86%). Only 6.67% and 3.47% of variation occurred within populations and among populations within lineages, respectively (Table 3), suggesting strong genetic divergence among the different regions. Additionally, a significant pattern of isolation-by-distance was detected across all studied locations using the Mantel test (*r*=0.873 1, *P*=0.001).

Table 3 AMOVA results of *T. heterochaetus* based on mtDNA

Source of variation	*d. f.*	Sum of squares	Variance component	Percentage of variation (%)	Fixation indices
Among lineages	2	3171.121	14.871 Va	89.86	F_{SC} =0.342*
Among populations within lineages	5	120.279	0.573 Vb	3.47	F_{ST} =0.933*
Within populations	312	344.225	1.103 Vc	6.67	F_{CT} =0.818*
Total	319	3635.625	16.548		

*: *P*<0.05; *d. f.*: Degree of freedom; Va: Variance component due to differences among lineages, Vb: Variance component due to differences among populations within lineages, Vc: Variance omponent due to differences among individuals within populations; F_{SC}=Vb/(Vb+Vc), F_{ST}=(Va+Vb)/(Va+Vb+Vc), F_{CT} =Va/(Va+Vb+Vc).

Historical demography

As genetic differences were not significant in the two geographically close location pairs FA-FZ and HP-ND, we treated them as a whole when analyzing demographic history.

Both Tajima's *D* and Fu's *F* tests showed negative values for all genetically differentiated populations (Table 5), indicating departure from mutation-drift equilibrium and possible population demographic expansion. Furthermore, mismatch
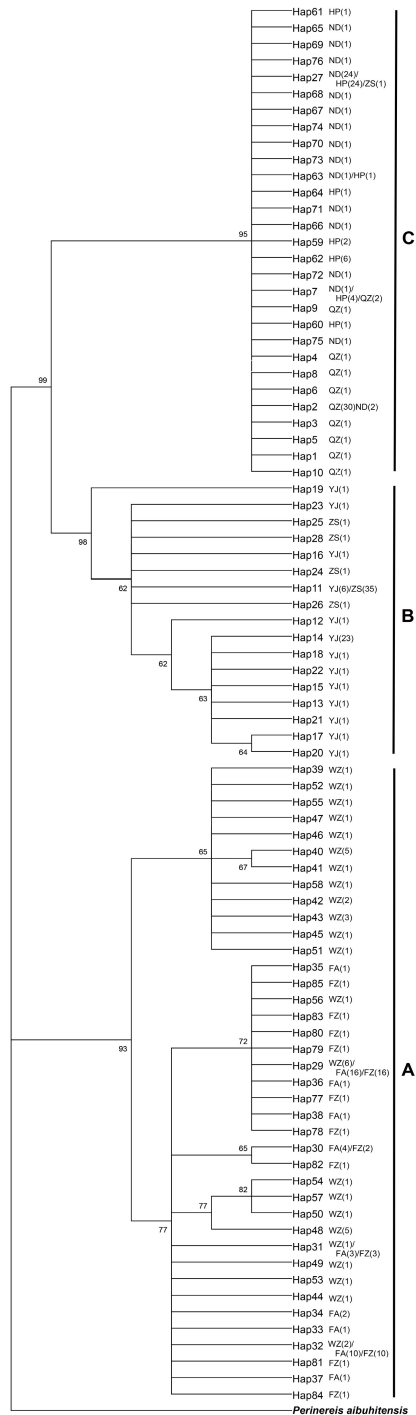
Figure 2 Neighbor-joining phylogenetic tree of *T. heterochaetus* based on *COI* haplotypes

Bootstrap values >60% are shown near nodes (1 000 replicates), number in parentheses after location code is frequency of haplotypes. *Perinereis aibuhitensis* (GenBank accession No.: NC023943.1) was used as an outgroup. WZ: Wenzhou; FA: Fu'an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dịnh.

Table 4 Pairwise Φ_{ST} (below diagonal) and genetic distance (above diagonal) among different *T. heterochaetus* locations

Location code	WZ	FA	FZ	ZS	YJ	QZ	HP	ND
WZ	–	0.009	0.009	0.058	0.057	0.061	0.061	0.060
FA	**0.226**	–	0.005	0.059	0.058	0.062	0.063	0.063
FZ	**0.238**	−0.013	–	0.059	0.058	0.062	0.063	0.063
ZS	**0.906**	**0.946**	**0.946**	–	0.002	0.020	0.019	0.019
YJ	**0.903**	**0.944**	**0.944**	**0.378**	–	0.022	0.021	0.021
QZ	**0.912**	**0.950**	**0.951**	**0.941**	**0.942**	–	0.005	0.005
HP	**0.907**	**0.944**	**0.947**	**0.922**	**0.926**	**0.711**	–	0.001
ND	**0.911**	**0.951**	**0.951**	**0.936**	**0.938**	**0.771**	0.017	–

Φ_{ST} values in bold type indicate statistical significance ($P<0.01$). WZ: Wenzhou; FA: Fu'an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dịnh.

distribution did not differ significantly from the model of sudden expansion when using either *SSD* or *HRI* for goodness-of-fit (Table 5), further supporting the hypothesis of population expansion in *T. heterochaetus*. The *τ* value across populations varied from 0.969 to 13.936. Using the mutation rate of 2% per million years, it was estimated that the *T. heterochaetus* population expansion occurred about 17 000–246 000 years ago in the middle to late Pleistocene.

DISCUSSION

Genetic diversity
This study examined the population genetic variability of *T. heterochaetus* using mtDNA for the first time. Compared with other polychaetes, the northern location (FA, FZ, WZ) in lineage A showed high haplotype and nucleotide diversities (Hd=0.861, π=0.007 11; Figure 1, Table 1), differing from several other species, such as *Branchipolynoe symmytilida* (Hd=0.970, π=0.007; Plouviez et al., 2009), *Owenia fusiformis* clade 1 and 2 (Hd=0.924–0.978, π=0.007 2–0.007 7; Jolly et al., 2006), and *Aglaophamus australiensis* (Hd=0.78–0.99; Smith et al., 2015). However, remarkably lower nucleotide diversity was detected in the southern locations (YJ, ZS, QZ, HP, and ND; Figure 1, Table 1), which may be due to the commercial overexploitation of wild stocks as *T. heterochaetus* is a very popular specialty food in this region. Although small-scale artificial breeding of *T. heterochaetus* has been conducted in southern China and Vietnam, production still comes primarily from natural exploitation.

Previous meta-analysis revealed that overharvesting can drive the decay of genetic diversity in many highly abundant marine species (Pinsky & Palumbi, 2014). Genetic diversity is closely related to the long-term adaptability and survivability of populations, especially in suddenly and drastically changing marine environments (Barrett & Schluter, 2008). Thus, the strikingly varied genetic diversity among different *T. heterochaetus* locations, as shown in this study, should be an important consideration in relation to conservation strategies or artificial breeding programs for future aquaculture enhancement.
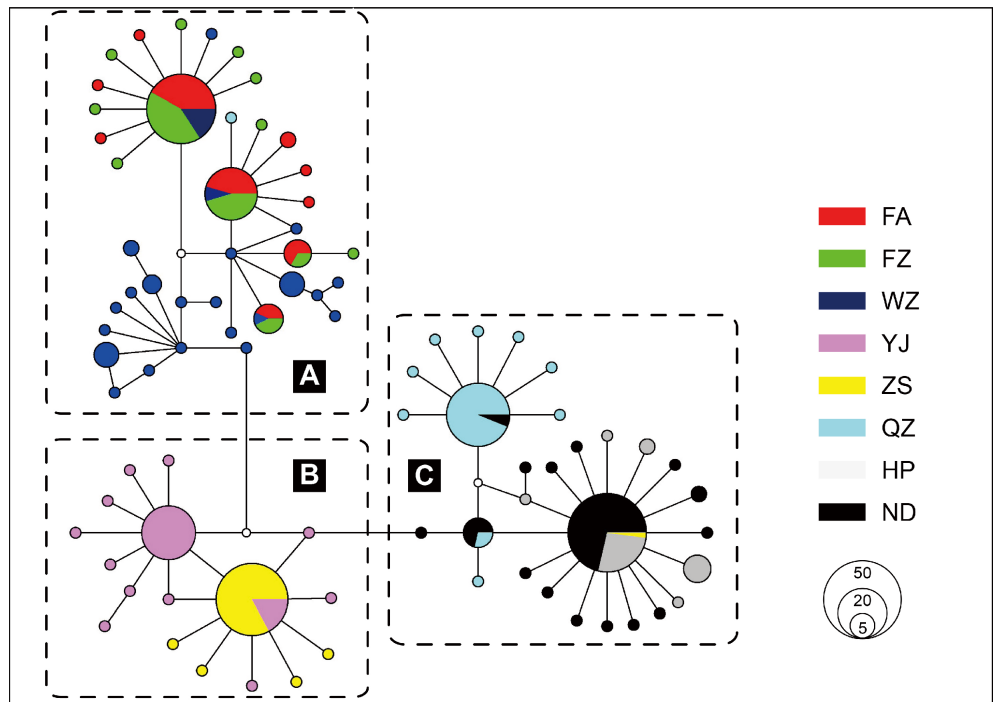
Figure 3 Median-joining network of *T. heterochaetus* based on *COI* haplotypes

Each circle represents a haplotype and area is proportional to its frequency. Colors of circles denote geographic origin, white dots represent hypothetical intermediate haplotypes. WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dịnh.

Table 5 Parameters of neutrality test and mismatch distribution analysis for six *T. heterochaetus* populations

Population	Neutrality test		Mismatch distribution analysis			Expansion time
	Tajima's *D*	Fu's *Fs*	*SSD*	*HRI*	Tau (*τ*)	*t* (Ma)
WZ	−0.486	−7.786**	0.023	0.024	6.425	0.113
FA+FZ	−0.395	−5.076*	0.052	0.123	4.834	0.085
ZS	−2.571**	−1.607	0.002	0.408	13.936	0.246
YJ	−1.995**	−11.090**	0.001	0.06	1.062	0.019
QZ	−2.172**	−7.960**	0.000 1	0.124	1.103	0.019
HP+ND	−2.333**	−24.297**	0.003	0.081	0.969	0.017

*: *P*<0.05, **: *P*<0.01. WZ: Wenzhou; FA: Fu' an; FZ: Fuzhou; ZS: Zhongshan; YJ: Yangjiang; QZ: Qinzhou; HP: Hai Phong; ND: Nam Dịnh.

Population structure and phylogenetic analysis

According to the pairwise Φ_{ST} values (Table 4), we identified six genetically different populations (WZ, FA+FZ, ZS, YJ, QZ, HP+ND) among the eight locations in this study. Given the geographic positions and significant divergences among the populations, Leizhou Peninsula and Hainan Island appear to be remarkable oceanographic barriers that block gene flow among the different populations inside the Beibu Gulf and others (Figure 1), as also demonstrated in *Perinereis aibuhitensis* populations from the southern coastal zone of China (Liu et al., 2014). Over broader geographic scales, long distance/isolation between populations may play an important role in the high level of genetic structuring. In fact, an evident isolation by distance pattern was detected using the Mantel test (*r*=0.873 1, *P*=0.001). As for the genetic homogeneity in

the two location pairs FA-FZ and HP-ND, the short distance along the coastline (<60 km, measured by Google Earth) may not be strong enough to prevent effective gene flow between the two adjacent estuaries.

Geographically, the southern populations (YJ, ZS, QZ, HP+ND) are isolated from the northern populations (FA+FZ, WZ) by several main water systems, such as the Pearl River, Hanjiang River, and Minjiang River. Such isolation could be a strong barrier to gene flow. In addition, according to our resource investigations over the last five years, many unsuitable habitat patches exist along the coastline among the sampling sites. This habitat discontinuity could also lead to the significant genetic divergences observed in different populations.

Apart from geographic factors, its narrow habitat niche is

also suspected to be partly responsible for the high structuring among *T. heterochaetus* populations. The pairwise Φ_{ST} results among sampling sites suggest striking genetic divergence, except for the FA-FZ and HP-ND pairs, which exhibit close geographical positions (Figure 1, Table 4), implying that effective dispersal of *T. heterochaetus* over large distances is limited. In marine invertebrates with sedentary adults, a longer planktonic larval stage is correlated with increased dispersal capacity and higher connectivity among populations (Cowen & Sponaugle, 2009; Kyle & Boulding, 2000). However, pelagic larva duration may not always be a good predictor of gene flow and population structure (Weersing & Toonen, 2009). Although it has a relatively long planktonic larval stage (~20 d), wild *T. heterochaetus* individuals favor the muddy bottoms of brackish environments with low salinity, such as estuaries (Tuan, 2018). This specialized habitat and narrow salinity requirement may limit colonization potential of the dispersed larvae. A high level of genetic divergence among invertebrates inhabiting estuarine systems is not uncommon (e. g., Darling et al., 2004; Olson et al., 2009; Virgilio et al., 2006), as estuaries represent spatially discrete habitats that tend to restrict gene flow and lead to different levels of isolation (Bilton et al., 2002).

In the present study, the phylogenetic tree, haplotype network, and AMOVA results indicated that the eight geographical locations of *T. heterochaetus* along the coasts of southeast China and north Vietnam could be characterized into three distinguishable lineages (i.e., A, B, C; Figures 2–3), corresponding to different geographic regions (A: East China Sea; B: Eastern Leizhou Peninsula; C: Western Leizhou Peninsula). Pairwise genetic distances among the 85 haplotypes ranged from 0.001 to 0.067, some of which were higher than 0.02, which is a commonly used standard for species identification (Avise, 2000). Hebert et al. (2003) suggested that higher intraspecific divergences ordinarily occur as geographic isolates, reflecting gene pool fragmentation in the origin of species in past episodes. In the last glacial maximum of the late-Pleistocene, the mean sea level was about 120 m lower than the current level (Fairbanks, 1989). During this period, the East China Sea shrank to a long, narrow ocean trough (Okinawa trough) and was isolated with the South China Sea (Koizumi et al., 2006). In addition, Hainan Island and the Chinese mainland were connected by the Qiongzhou Strait land-bridge (Voris, 2000). These historical geographic barriers may have blocked gene flow among different coastal regions and led to the deep lineages observed in the current *T. heterochaetus* populations.

Population genetic analysis can provide guidelines for strategies in species conservation and germplasm resource management (Loeschcke et al., 2013; Whiteley et al., 2006). Different lineages can show varying degrees of physiological adaptations (Méndez et al., 2001). Thus, we suggest that the genetically divergent lineages confirmed in this study should be treated as independent management units regarding conservation issues. Based on the relatively low level of genetic diversity in samples from ZS and QZ, special attention

should be paid to these two populations. As a typical *r*-selected species, *T. heterochaetus* is characterized by its short life span (one year) and hypersensitivity to environmental deterioration, habitat fragmentation, and overfishing (Tuan, 2018). To effectively protect the local populations of *T. heterochaetus*, water pollution control, overfishing reduction, and habitat preservation are all high priorities for future conservation.

Previous studies have suggested that the genetic structure of marine invertebrates can be shaped by various factors, such as currents, geographic segregation, life history characteristics, human-mediated transfer, and local selection (Selkoe & Toonen, 2011; Simon & Sato-Okoshi, 2015; Zakas & Wares, 2012; Zardi et al., 2007). Therefore, further research is still needed to identify the underlying mechanisms that may contribute to the substantial genetic heterogeneity in *T. heterochaetus*.

Historical demography

Both mismatch distribution analysis and significant negative values of neutrality tests (Tajima's *D* and Fu's *Fs*) indicate a pattern of recent population expansion in *T. heterochaetus*. This hypothesis was also supported by the star-shaped haplotype network (Figure 3), a characteristic of exponential population growth (Slatkin & Hudson, 1991). The demographic history was reflected in the genetic indices of this species, which showed low nucleotide diversity (π=0.032 15) but high haplotype diversity (*Hd*=0.926) (Table 2). As described by Grant & Bowen (1998), high *Hd* and low π can be attributed to rapid population expansion that enhances the retention of new mutations, which is consistent with the large number of unique and low-frequency haplotypes found within lineages in the present study (Figure 3).

Due to the lack of fossil and geological records, which are major obstacles for phylogeographic analysis of marine invertebrates (Provan & Bennett, 2008), the precise population expansion time based on a species-specific molecular clock for *T. heterochaetus* is not available. Based on a mutation rate of 2% per million years, our rough estimation suggests population expansion for *T. heterochaetus* in the middle to late Pleistocene, a period dominated by glaciation cycles (Imbrie et al., 1992) and periodic climatic oscillations, which may have impacted the distribution of *T. heterochaetus*. A similar demographic history pattern has also been found in other marine invertebrates along the Chinese coast (Gao et al., 2018; Liu et al., 2012).

CONCLUSIONS

Based on partial sequences of the *COI* gene, this study provides preliminary information on the genetic status of *T. heterochaetus* collected from the coasts of southeast China and north Vietnam. The genetic diversity in this species was highly variable and we identified six populations with significant genetic differences. These six populations could be divided into three genetically divergent lineages, corresponding to three geographic regions. We suggest that

the lineages confirmed in this study should be managed separately from a conservation point of view. For future studies, the application of multiple markers with higher resolution (e. g., microsatellites and single-nucleotide polymorphisms) and greater spatial sampling, as well as a broader understanding of biological and ecological factors, should provide a more detailed assessment of the *T. heterochaetus* population structure.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Z.N.M. and S.Y. conceived and designed the experiment; X.H.C. and S.Y. performed the experiments; X.H.C. and Z.N.M. wrote the manuscript; W.Y., Y.Y.S., and R.W.X. contributed reagents/materials/analysis tools; B.F. and L. W. collected the samples and made English corrections. All authors read and approved the final version of the manuscript.

REFERENCES

Avise JC. 2000. Phylogeography: the History and Formation of Species. Cambridge, MA: Harvard University Press.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**(1): 37−48.

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**(1): 38−44.

Bilton DT, Paula J, Bishop JDD. 2002. Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal and Shelf Science*, **55**(6): 937−952.

Birky CW, Fuerst P, Maruyama T. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**(3): 613−627.

Blank M, Bastrop R. 2009. Phylogeny of the mud worm genus *Marenzelleria* (Polychaeta, Spionidae) inferred from mitochondrial DNA sequences. *Zoologica Scripta*, **38**(3): 313−321.

Chen X, Li M, Liu H, Li B, Guo L, Meng Z, Lin H. 2016. Mitochondrial genome of the polychaete *Tylorrhynchus heterochaetus* (Phyllodocida, Nereididae). *Mitochondrial DNA Part A*, **27**(5): 3372−3373.

Costa PFE, Gil J, Passos A M, Pereira P, Melo P, Batista F, Da Fonseca LC. 2006. The market features of imported non-indigenous polychaetes in Portugal and consequent ecological concerns. *Scientia Marina*, **70**(S3): 287−292.

Cowen RK, Sponaugle S. 2009. Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, **1**: 443−466.

Darling JA, Reitzel AM, Finnerty JR. 2004. Regional population structure of a widely introduced estuarine invertebrate: *Nematostella vectensis* Stephenson in New England. *Molecular Ecology*, **13**(10): 2969−2981.

Dean HK. 2008. The use of polychaetes (Annelida) as indicator species of marine pollution: a review. *Revista de Biologia Tropical*, **56**(4): 11−38.

Duan XH, Liu WS, Xu AY, Xu CH, He QT, Xiao RY, Pan G. 2017. Observation on embryonic development of *Tylorrhynchus heterochaeta* and effect of salinity on their hatching. *South China Fisheries Science*, **13**(4): 115−121.

Excoffier L, Layal G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**: 47−50.

Fairbanks RG. 1989. A 17, 000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deepocean circulation. *Nature*, **342**(6250): 637−642.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**(5): 294−299.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**(2): 915−925.

Gao X, Mu D, Lou J, Hou C, Zhu J, Wang J, Jin S. 2018. Genetic diversity and structure analysis of *Phascolosoma esculenta* in the coastal zone of south-eastern China based on mitochondrial Cyt b gene. *Turkish Journal of Fisheries and Aquatic Sciences*, **18**(4): 519−528.

Glasby CJ, Timm T. 2008. Global diversity of polychaetes (Polychaeta; Annelida) in freshwater. *Hydrobiologia*, **595**: 107−115.

Grant WAS, Bowen BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**(5): 415−426.

Green BN, Suzuki T, Gotoh T, Kuchumov AR, Vinogradov SN. 1995. Electrospray ionization mass spectrometric determination of the complete polypeptide chain composition of *Tylorrhynchus heterochaetus* hemoglobin. *Journal of Biological Chemistry*, **270**(31): 18209−18211.

Harpending H. 1994. Signature of ancient population growth in a lowresolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**(4): 591−600.

Hebert PDN, Ratnasingham S, de Waard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**(S1): S96−S99.

Imbrie J, Boyle EA, Clemens SC, Duffy A, Howard WR, Kukla G, Kutzbach J, Martinson DG, McIntyre A, Mix AC, Molfino B, Morley JJ, Peterson LC, Pisias NG, Prell WL, Raymo ME, Shackleton NJ, Toggweiler JR. 1992. On the structure and origin of major glaciation cycles 1. Linear responses to Milankovitch forcing. *Paleoceanography and Paleoceanography*, **7**(6): 701−738.

Jolly MT, Viard F, Gentil F, Thiébaut É, Jollivet D. 2006. Comparative phylogeography of two coastal polychaete tubeworms in the Northeast Atlantic supports shared history and vicariant events. *Molecular Ecology*, **15**(7): 1841−1855.

Kaoawa Y. 1954. Histological Observations on the transformation to mature worms of *Tylorrhynchus heterochaetus*. *Journal of Gakugei, Tokushima University*, **4**: 65−71.

Kesäniemi JE, Rawson PD, Lindsay SM, Knott KE. 2012. Phylogenetic analysis of cryptic speciation in the polychaete *Pygospio elegans*. *Ecology and Evolution*, **2**(5): 994−1007.

Koizumi I, Tada R, Narita H, Irino T, Aramaki T, Oba T, Yamamoto H. 2006. Paleoceanographic history around the Tsugaru Strait between the Japan Sea and the Northwest Pacific Ocean since 30 cal kyr BP. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **232**(1): 36−52.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870−1874.

Kyle CJ, Boulding EG. 2000. Comparative population genetic structure of marine gastropods (*Littorina spp.*) with and without pelagic larval dispersal. *Marine Biology*, **137**(5-6): 835−845.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**(11): 1451−1452.

Liu F, Guo QS, Shi HZ, Lv F, Yu YB, Lv LL. Huang JT, Wang AM, Liang HX. 2014. Genetic diversity analysis of *Perinereis aibuhitensis* based on ISSR and SRAP markers of Chinese coast populations. *Biochemical Systematics and Ecology*, **57**: 262−269.

Liu H, Liu M, Ge S, Wang Q, Yu D, Guan S. 2012. Population structuring and historical demography of a common clam worm *Perinereris aibuhitensis*, near the coasts of Shandong Peninsula. *Biochemical Systematics and Ecology*, **44**(4): 70−78.

Loeschcke V, Tomiuk J, Jain SK. 2013. Conservation genetics (Vol. 68. Basel: Birkhäuser, 37-53.

Méndez N, Linke-Gamenick I, Forbes VE, Baird DJ. 2001. Sediment processing in *Capitella spp.* (Polychaeta: Capitellidae): strain-specific differences and effects of the organic toxicant fluoranthene. *Marine Biology*, **138**(2): 311−319.

Okada K. 1952. Experimental studies on the Japanese palolo, *Tylorrhynchus heterochaetus. I. activation of egg-cells. Journal of Gakugei, Tokushima University*, **2**: 43−52.

Olson MA, Zajac RN, Russello MA. 2009. Estuarine-scale genetic variation in the polychaete *Hobsonia florida* (Ampharetidae; Annelida) in Long Island Sound and relationships to Pleistocene glaciations. *The Biological Bulletin*, **217**(1): 86−94.

Pinsky ML, Palumbi SR. 2014. Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology*, **23**(1): 29−39.

Plouviez S, Shank TM, Faure B, Daguin-Thiebaut C, Viard F, Lallier FH, Jollivet D. 2009. Comparative phylogeography among hydrothermal vent species along the East Pacific Rise reveals vicariant processes and population expansion in the South. *Molecular Ecology*, **18**(18): 3903−3917.

Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**(10): 564−571.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**(3): 552−569.

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**(4): 406−425.

Sato M, Osanai K. 1990. Sperm attachment and acrosome reaction on the egg surface of the polychaete, *Tylorrhynchus heterochaetus. The Biological Bulletin*, **178**(2): 101−110.

Selkoe KS, Toonen RJ. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, **436**: 291−305.

Simon CA, Sato-Okoshi W. 2015. Polydorid polychaetes on farmed molluscs: distribution, spread and factors contributing to their success. *Aquaculture Environment Interactions*, **7**(2): 147−166.

Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**(2): 555−562.

Smith LM, Hutchings P, Fraser CI. 2015. Molecular evidence supports coastal dispersal among estuaries for two benthic marine worm (Nephtyidae) species in southeastern Australia. *Marine Biology*, **162**(6): 1319−1327.

Suzuki T, Gotoh T. 1986. The complete amino acid sequence of giant multisubunit hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. *Journal of Biological Chemistry*, **261**(20): 9257−9267.

Suzuki T, Takagi T, Gotoh T. 1990. Primary structure of two linker chains of the extracellular hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. *Journal of Biological Chemistry*, **265**(21): 12168−12177.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**(3): 585−595.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**(24): 4876−4882.

Tuan NN. 2018. Biological characteristics and effects of salinity on reproductive activities of marine worm (*Tylorrhynchus heterochaetus*, Quatefages, 1865) in summer season in Hai Phong - Viet Nam. *SNRU Journal of Science and Technology*, **10**(1): 25−31.

Virgilio M, Backeljau T, Abbiati M. 2006. Mitochondrial DNA and allozyme patterns of Hediste diversicolor (Polychaeta: Nereididae): the importance of small scale genetic structuring. *Marine Ecology Progress Series*, **326**: 157−165.

Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**(5): 1153−1167.

Weersing K, Toonen RJ. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**: 1−12.

Whiteley AR, Spruell P, Allendorf FW. 2006. Can common species provide valuable information for conservation?. *Molecular Ecology*, **15**(10): 2767−2786.

Zakas C, Wares JP. 2012. Consequences of a poecilogonous life history for genetic structure in coastal populations of the polychaete *Streblospio benedicti*. *Molecular Ecology*, **21**(22): 5447−5460.

Zardi GI, McQuaid CD, Teske PR, Teske PR, Barker NP. 2007. Unexpected genetic structure of mussel populations in South Africa: indigenous *Perna perna* and invasive *Mytilus galloprovincialis*. *Marine Ecology Progress Series*, **337**: 135−144.

Tube-nosed variations–a new species of the genus *Murina* (Chiroptera: Vespertilionidae) from China

DEAR EDITOR,

During a survey in 2014, several tube-nosed bats (Vespertilionidae: Murininae: *Murina*) were collected in Sichuan Province. Based on morphological characters, these bats did not match any species previously recorded from China. Morphometric analyses and phylogenetic inference based on mitochondrial and nuclear gene sequences indicated that they represented a new species, named here as *Murina jinchui* **sp. nov**. Although the new species is presently known only from Wolong National Nature Reserve, it is unlikely to be a rare species in the area based on our capture frequencies.

Characterized by tubular nostrils and relatively well-developed anterior upper premolars, the Old World subfamily of vespertilionid bats, Murininae Miller, 1907, is rich in cryptic species. Typically, these species are rare in collections, which have contributed to our poor understanding of their diversity and distribution. Simmons (2005) listed 17 species within the subfamily, but several new species have been described since due to an increase in survey efforts, improved capture methods, re-evaluation of taxonomically informative characters, and species delimitations using DNA barcoding (Csorba et al., 2011; Eger & Lim, 2011; Francis et al., 2010). As such, 39 species are currently recognized based on taxonomic revisions and new species descriptions (Chen et al., 2017; Csorba et al., 2007, 2011 Eger & Lim, 2011; Francis & Eger, 2012; Furey et al., 2009; He et al., 2015; Kruskop & Eger, 2008; Kuo et al., 2009; Maeda & Matsumura, 1998; Ruedi et al., 2012; Soisook et al., 2013a, 2013b Son et al., 2015; Zeng et al., 2018). In the last decade, intensive survey efforts and morphological and molecular studies have resulted in the description of nine new *Murina* species from China alone, namely *M. bicolor*, *M. gracilis*, and *M. recondita* from Taiwan (Kuo et al., 2009), *M. chrysochaetes*, *M. lorelieae*, and

M. shuipuensis from Guangxi and Guizhou (Eger & Lim, 2011), and *M. fanjingshanensis* (He et al., 2015), *M. rongjiangensis* (Chen et al., 2017), and *M. liboensis* (Zeng et al., 2018) from Guizhou. Thus, at least 19 species belonging to the genus *Murina* are currently known from China (Chen et al., 2017; Eger & Lim, 2011; He et al., 2015; Jiang et al., 2015; Kuo et al., 2009; Liu & Wu, 2019; Zeng et al., 2018).

In 2014, several small-sized and impressively colored *Murina* individuals were collected during field surveys in Sichuan Province (all field surveys and sample collection protocols complied with the current laws of Sichuan Province). Morphological and molecular biological examinations revealed them to be distinct from all other recognized *Murina* taxa; therefore, they are described herein as a new species.

For morphometric analysis, we examined 224 specimens of *Murina* deposited in 12 collections (see list of specimens in Supplementary Material (Appendix I)). Six external measurements (to the nearest 0.1 mm), body mass (to the nearest 0.1 g), and 16 craniodental measurements (to the nearest 0.01 mm) were recorded by the same author. Definitions and details of measurements are listed in the Table 1 and Supplementary Material (Methods).

We performed principal component analysis (PCA) and discriminant analysis of principal components (DAPC) (Jombart, 2008; Jombart et al., 2010) for species discrimination based on external and craniodental measurements. For both PCA and DAPC, the sexes were analysed separately because sexual dimorphism is noted in

Open Access

Table 1 Selected external and craniodental measurements (mm) of *Murina jinchui* sp. nov and four closely related *Murina* species

Item	*Murina jinchui* **sp. nov.**			*M. rongjiangensis*		*M. shuipuensis*		*M. fanjing-shanensis*	*M. leucogaster*	
	♀♀	♂♂	*t*-value	♀♀	♂♂	♀♀	♂♂	♂	♀♀	♂
GTL	16.47 (3) (16.10–16.86)	15.73 (3) (15.66–15.83)	3.25*	16.27±0.3 (5) (15.88–16.71)	15.76±0.29 (5) (15.32–16.04)	16.00, 16.05 (2)	15.82 (3) (15.60–16.15)	19.22 (1)	19.14 (3) (18.77–19.33)	19.5 (1)
CCL	14.42 (3) (14.34–14.48)	13.65 (3) (13.58–13.68)	14.5*	14.38±0.45 (5) (13.73–14.98)	13.84±0.21 (5) (13.49–14.03)	13.92, 14.26 (2)	13.90 (3) (13.82–14.00)	16.13 (1)	16.81 (3) (16.52–17.13)	16.84 (1)
CBL	15.81 (3) (15.50–16.07)	15.14 (3) (15.12–15.15)		15.19±0.48 (5) (14.51–15.78)	14.74±0.14 (5) (14.60–14.95)	14.65, 15.06 (2)	14.64 (3) (14.53–14.83)	17.62 (1)	17.85 (3) (17.60–18.09)	17.91 (1)
BCW	7.64 (3) (7.60–7.72)	7.18 (3) (7.15–7.22)	10.18*	7.63 ±0.21 (5) (7.45–7.94)	7.57±0.1 (5) (7.44–7.71)	7.31, 7.47 (2)	7.33 (3) (7.14–7.42)	9.43 (1)	8.77 (3) (8.17–9.10)	8.77 (1)
BCH	7.54 (3) (7.43–7.69)	7.26 (3) (7.04–7.41)	2.05NS	7.4 ±0.25 (5) (7.00–7.64)	7.2±0.15 (5) (7.00–7.37)	6.98, 7.13 (2)	6.96 (3) (6.78–7.22)	8.74 (1)	8.27 (3) (8.13–8.34)	8.16 (1)
ZYW	8.64 (3) (8.53–8.79)	8.55 (3) (8.49–8.60)	0.88NS	9.08 ±0.26 (5) (8.66–9.37)	8.67±0.26 (5) (8.47–9.06)	8.58, 8.70 (2)	8.32 (3) (8.06–8.60)	9.22 (1)	10.74 (3) (10.48–10.99)	10.84 (1)
MAW	8.01 (3) (7.94–8.10)	7.64 (3) (7.54–7.83)	3.53*	7.97 ±0.2 (5) (7.69–8.19)	7.74±0.16 (5) (7.60–7.99)	7.56, 7.68 (2)	7.47 (3) (7.37–7.61)	8.66 (1)	9.10 (3) (8.85–9.21)	9.31 (1)
PL	7.31 (3) (7.09–7.58)	7.15 (3) (7.03–7.33)	0.92NS	7.00 ±0.16 (5) (6.79–7.19)	6.8±0.16 (5) (6.70–7.08)	7.50, 7.56 (2)	7.47 (3) (7.13–7.69)	9.22 (1)	8.66 (3) (8.44–8.97)	8.74 (1)
IOW	4.41 (3) (4.32–4.52)	4.24 (3) (4.05–4.50)	1.15NS	4.69 ±0.13 (5) (4.54–4.90)	4.54±0.15 (5) (4.30–4.70)	4.33, 4.57 (2)	4.47 (3) (4.30–4.65)	5.46 (1)	5.39 (3) (5.10–5.69)	5.50 (1)
CM³L	5.52 (3) (5.47–5.58)	5.21 (3) (5.12–5.32)	4.68*	5.31 ±0.14 (5) (5.13–5.47)	5.06±0.06 (5) (4.97–5.13)	5.28, 5.33 (2)	5.33 (3) (5.25–5.39)	6.35 (1)	6.46 (3) (6.36–6.53)	6.34 (1)
CCW	4.05 (3) (3.99–4.16)	3.79 (3) (3.65–3.88)	2.89*	4.26 ±0.15 (5) (4.02–4.38)	3.92±0.14 (5) (3.80–4.13)	3.94, 4.08 (2)	3.92 (3) (3.90–3.94)	4.70 (1)	4.74 (3) (4.62–4.91)	4.91 (1)
M³M³W	5.92 (3) (5.88–5.95)	5.44 (3) (5.32–5.51)	7.28*	5.71 ±0.18 (5) (5.49–5.95)	5.53±0.19 (5) (5.27–5.76)	5.40, 5.57 (2)	5.49 (3) (5.38–5.54)	6.54 (1)	6.51 (3) (6.04–6.79)	6.66 (1)
RCM	0.68 (3) (0.67–0.71)	0.70 (3) (0.69–0.70)	−0.97NS	0.75 ±0.02 (5) (0.73–0.78)	0.71±0.02 (5) (0.69–0.73)	0.73, 0.73 (2)	0.71 (3) (0.70–0.73)	0.72 (1)	0.73 (3) (0.69–0.81)	0.74 (1)
LCM₃L	5.78 (3) (5.68–5.85)	5.68 (3) (5.57–5.75)	1.32NS	5.59 ±0.11 (5) (5.44–5.71)	5.52±0.11 (5) (5.42–5.68)	5.57, 5.61 (2)	5.65 (3) (5.53–5.75)	–	6.73 (3) (6.49–6.85)	6.58 (1)
ML	11.41 (3) (11.34–11.49)	10.79 (3) (10.68–10.90)	8.08*	11.25 ±0.32 (5) (10.98–11.78)	10.65±0.23 (5) (10.38–11.00)	10.84, 10.98 (2)	10.75 (3) (10.63–10.83)	–	12.99 (3) (12.71–13.22)	12.98 (1)
CPH	3.95 (3) (3.90–4.01)	3.37 (3) (3.34–3.42)	14.21*	3.8 ±0.19 (5) (3.49–3.99)	3.48±0.23 (5) (3.24–3.80)	3.73, 3.88 (2)	3.51 (3) (3.39–3.68)	–	4.86 (3) (4.62–5.02)	4.97 (1)
Wt	7.3 (3) (7.1–7.5)	4.9 (3) (4.5–5.2)		6.8±1.09 (5) (5.0–7.8)	5.0±0.26 (5) (4.7–5.4)	4.3, 5.1 (2)	4.7 (3) (4.3–5.0)	9.0 (1)	–	10.9 (1)
HB	44.7 (3) (43.0–45.7)	40.1 (3) (39.7–40.4)	5.45*	43.5 ± 2.48 (5) (40.6–46.7)	38.8 ± 4.86 (5) (34.0–46.5)	40.6, 40.9 (2)	41.4 (3) (40.0–43.3)	46.3 (1)	47.7 (3) (44.4–51.4)	50.3 (1)
T	34.9 (3) (33.0–38.4)	34.9 (3) (33.6–36.0)	0.01NS	34.9 ± 1.90 (5) (32.9–37.3)	33.8 ± 1.38 (5) (32.3–35.4)	32.0, 38.4 (2)	33.5 (3) (31.9–36.6)	39.7 (1)	40.1 (3) (37.7–41.5)	39.3 (1)
E	15.3 (3) (14.4–16.1)	14.7 (3) (14.4–15.1)	1.08NS	11.2 ± 0.72 (5) (10.2–11.9)	10.6 ± 0.98 (5) (9.0–11.6)	11.7, 13.4 (2)	11.6 (3) (10.7–12.4)	15.3 (1)	12.8 (3) (12.4–13.3)	13.8 (1)
HF	8.2 (3) (7.2–9.4)	7.5 (3) (7.4–7.7)	1.03NS	7.8 1.07 (5) (5.9–8.7)	7.5 ± 0.89 (5) (6.1–8.4)	8.7, 9.1 (2)	8.5 (3) (7.8–9.1)	8.5 (1)	9.6 (3) (8.9–10.3)	11.4 (1)
FA	36.3 (3) (36.1–36.4)	33.5 (3) (32.4–34.6)	4.26*	33.0 ±1.36 (5) (30.6–34.0)	31.6 ± 0.81 (5) (31.0–32.7)	30.0, 32.7 (2)	30.8 (3) (29.7–31.9)	42.6 (1)	43.6 (3) (43.0–44.2)	41.5 (1)

Item	Murina jinchui sp. nov.			M. rongjiangensis		M. shuipuensis		M. fanjing-shanensis	M. leucogaster	
	♀♀	♂♂	t-value	♀♀	♂♂	♀♀	♂♂	♂	♀♀	♂
Tib	16.4 (3) (16.4–16.5)	15.3 (3) (14.7–16.0)	3.01[*]	14.8 ± 0.72 (5) (13.9–15.6)	15.0 ± 0.63 (5) (14.2–15.7)	14.5, 15.2 (2)	14.7 (3) (13.8–16.2)	19.5 (1)	19.3 (3) (18.3–19.9)	17.9 (1)

Values are given as Means±SD (if $n>5$), and minimum-maximum (min-max). t-value is from Students t-test between sexes when distribution of measurement fits normality, and [*] and [NS] represent $P<0.05$ and non-significant result, respectively. Abbreviations and definitions for measurements are as followed: HB: Total length–from the tip of the face/chin to the anus; T: Tail vertebrae length–from the tip of the tail to the beginning of the tail vertebrae; E: Ear length–from the notch at the base of the ear conch to the tip of the pinna; HF: Hind foot length–from the heel to the tip of the longest toe, including the claw; Tib: Length of tibia–from the knee to the ankle; FA: Forearm length–from the elbow to the wrist with both joints folded; GTL: Greatest length of skull–from the posterior edge of the skull to the front of the incisors; CCL: Condylocanine length–from the exoccipital condyle to the most anterior part of the canine; CBL: Condylobasal length–from the exoccipital condyle to the posterior rim of alveolus of the first upper incisor; BCW: Braincase width–greatest width across the braincase; BCH: Braincase height–from the basisphenoid at the level of the hamular processes to the highest part of the skull, including the sagittal crest (if present); ZYW: Zygomatic width–the greatest width of the skull across the zygomatic arches; MAW: Mastoid width–the greatest distance across the mastoid region; PL: Palatal length–from the anterior palatal emargination to the midpoint of the posterior palatal emargination; IOW: Interorbital width–least width of the interorbital constriction; CM^3L: Length of maxillary toothrow–from the front of the canine to the posterior edge of the 3rd upper molar; CCW: Greatest breadth across the upper canines; M^3M^3W: Width across upper molars–greatest width measured across the outer edges of the second upper molars; RCM: Ratio of CCW to M^3M^3W; LCM_3L: Length of mandibular toothrow–from the front of the canine to the posterior edge of the 3rd lower molar; ML: Greatest length of mandible–greatest length measured from the posterior edge of the mandibular condyles to the front of the lower incisors; CPH: Coronoid process height–measured from the inferior surface of the angular process of the ramus to the tip of the coronoid process.

several *Murina* species (Kuo et al., 2009; Son et al., 2015), including the new one (Table 1; Supplementary Figure S1 and Table S1). We also replicated our multivariate statistical analyses in the monophyletic clade formed by *M. leucogaster*, *M. shuipuensis*, *M. rongjiangensis*, and the new species. In addition, due to their similarities in size and skull proportions, *M. shuipuensis*, *M. rongjiangensis*, and the new species were also compared using analysis of variance (ANOVA). All analyses were performed using the "phych" and "adegenet" packages in R (Jombart, 2008; R Development Core Team, Vienna, www.R-project.org).

For phylogenetic analysis, the partial cytochrome oxidase subunit I (*COI*, 670 bp) and partial nuclear recombination activating protein 2 genes (*Rag2*, 1 339 bp) were selected as molecular markers (Heaney et al., 2012; Kuo et al., 2017; Lack & Bussche, 2010; Roehrs et al., 2010). The *COI* gene was amplified from specimens of all species in this study, whereas the *Rag2* gene was sequenced from single individuals of each species confirmed by our phylogenetic and morphological species determination (GenBank accession Nos.: MN549027–MN549101). All available *COI* and *Rag2* sequences of *Murina* collected from NCBI-nt and our specimens were aligned using MUSCLE (Edgar, 2004). Final alignment was partitioned by different codon positions and the parameters of the best nucleotide substitution models were determined by PartitionFinder2 (Lanfear et al., 2017) using the greedy algorithm (Lanfear et al., 2012). Maximum-likelihood (ML) trees were searched in RAxML v7.4.2 (Stamatakis et al., 2008), and the reliability of nodes was evaluated by 500 rapid bootstrap matrixes.

To describe pelage color, digital photographs of freshly euthanized bats were taken in the field (see details in Supplementary Material). Following Davis & Castleberry (2010), images were color calibrated, and pelage color was described following the Pantone color code with the red-green-blue (RGB) system.

Detailed measurements of all specimens used in this study are listed in Table 1. The PCA based on six external measurements revealed that 76% of total variance could be explained by the first two components (54% and 22% for principal components 1 (PC1) and principal components 2 (PC2), respectively) (Figure 1A). For PC1, all measurements, except for ear length, had high positive loadings, thus reflecting an external overall size effect. PC2 was mostly related to ear size (larger value indicates larger measurement) (Figure 1A, with values provided in Table 2). Although the interspecific patterns revealed by PCA were ambiguous, implying difficulty in species identification based on external measurements only, the PCA results consistently indicated larger females and smaller males within *Murina* species. For PCA of the 16 craniodental measurements, 81% of total variance could be explained by the first two components (57% and 24% for PC1 and PC2, respectively) (Figure 1B). For PC1, 11 measurements had positive loadings (Figure 1B, with values provided in Table 1), suggesting that this PC was mainly related to skull size (smaller bats characterized by lower PC1 scores). Thus, with the exception of *M. leucogaster*, most *Murina* species with *suilla*-type dentition (crown area of upper canine less than that of P^4) had smaller skulls than those species with *cyclotis*-type dental characters (crown area of upper canine equal to or larger than that of P^4) (Figure 1B). For PC2, most measurements had low loadings, except for braincase width (BCW), mastoid width (MAW), and interorbital width (IOW) (Figure 1B and Table 1). Unlike the ambiguous pattern found for external measurements, PCA based on craniodental measurements revealed a more

noticeable interspecific relationship, with six skull-size assemblages identified: i.e., (1) *M. leucogaster*; (2) *M. aurata*; (3) all taxa with *cyclotis*-type dentition included in this study; (4) *M. eleryi*+*M. chrysochaetes*; (5) *M. shuipuensis*+*M. rongjiangensis*+new species; and (6) "*M. jaintiana*+*M. beelzebub*+*M. feae*" (Figure 1B).
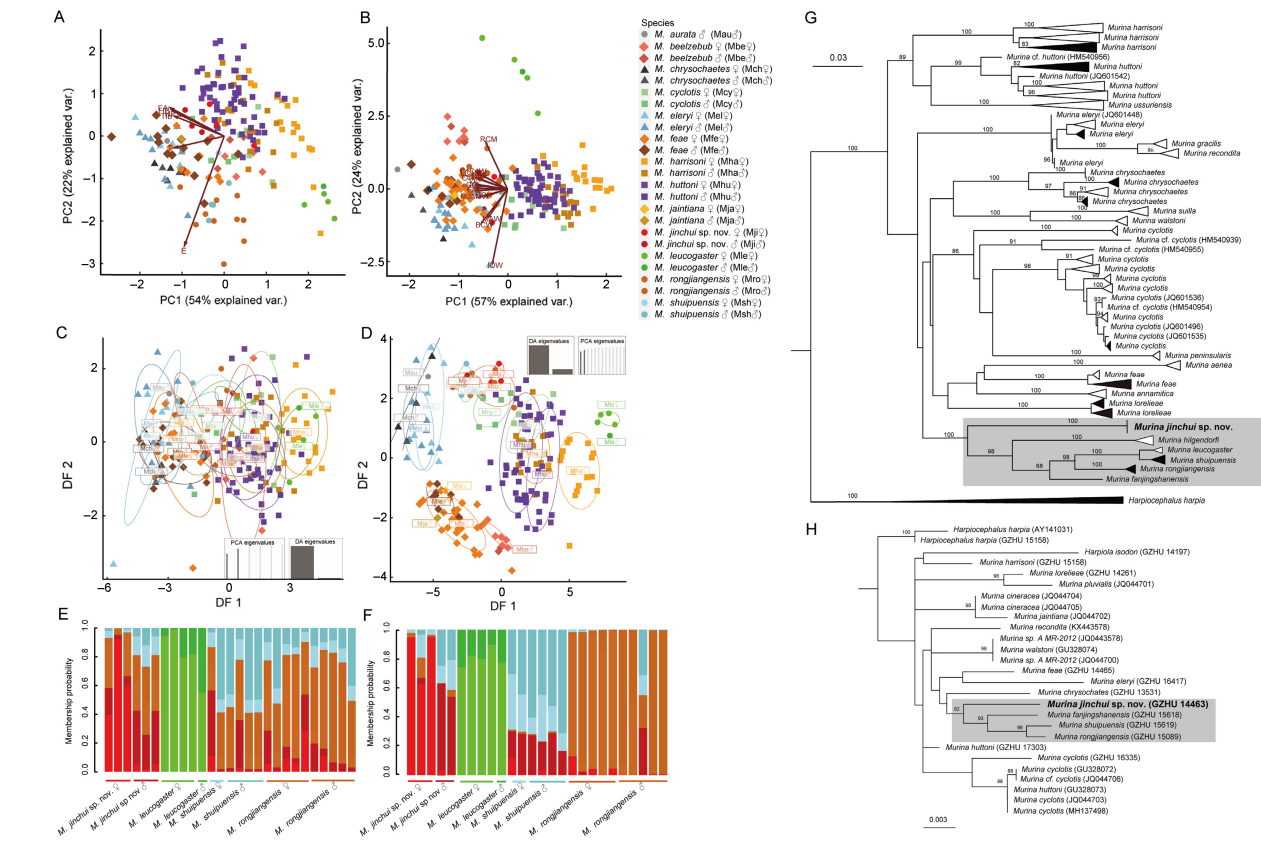

Figure 1 Two-dimensional PCA and DAPC plots and maximum-likelihood phylogenetic trees

A, B: PCA plots for *M. aurata*, *M. chrysochaetes*, *M. cyclotis*, *M. eleryi*, *M. harrisoni*, *M. huttoni*, *Murina jinchui* **sp. nov**., *M. leucogaster*, *M. rongjiangensis*, *M. shuipuensis*, *M. jaintiana*, *M. beelzebub*, and *M. feae* showing projections of individual specimens and variable loadings on first two principal components. C, D: Projections of 224 specimens and variable loadings on two DFs obtained from external and craniodental measurements. E, F: Posterior probabilities of the same 26 specimens reclassified following analysis of phylogenetic clade of *M. leucogaster*, *Murina jinchui* **sp. nov**., M. *shuipuensis*, and *M. rongjiangensis* using DAPC. According to species and sex, eight groups were predetermined in this DAPC. Maximum-likelihood trees based on mitochondrial *COI* (G) and nuclear *Rag2* (H) sequences for the subfamily Murininae. Triangles in (G) represent clusters of multiple specimens, with horizontal dimension proportional to amount of sequence divergence. Solid triangles indicate newly generated *COI* sequences. Numbers above branches indicate level of bootstrap support >50% for the branch.

We used the PC1 and PC2 values from the external and craniodental analyses for DAPC, with the results together explaining 87% and 90% of total variance, respectively. Among the original external variables, HB, T, and FA had high positive loadings on discriminant function 1 (DF1), whereas HB and T had high positive loadings on discriminant function 2 (DF2) (Table 2). For craniodental DAPC analysis, greatest length of skull (GTL), condylocanine length (CCL), and greatest length of mandible (ML) contributed substantially to DF1, whereas braincase height (BCH) and mastoid width (MAW) contributed substantially to DF2 (Table 2). Similar to the PCA plots of external variables, the DAPC plots revealed an ambiguous pattern, with most taxa overlapping (Figure 1C).

Nevertheless, based on an increase in discriminant power, the pattern of the craniodental DAPC plots was more distinguishable than that of the PCA plots, although most were still partially overlaid (Figure 1D). Similar to the PCA pattern from craniodental measurements, a noticeable interspecific relationship emerged, with five skull-size assemblages identified, including: (1) *M. leucogaster*; (2) all *cyclotis*-type taxa included in this study; (3) *M. aurata*+*M. eleryi*+*M. chrysochaetes*; (4) *M. shuipuensis*+*M. rongjiangensis*+new species; and (5) "*M. jaintiana*+*M. beelzebub*+*M. feae*" (Figure 1D). Based on DAPC reclassification of inference of clades, including *M. shuipuensis*, *M. rongjiangensis*, *M. leucogaster*, and the new species, craniodental DAPC revealed higher

Table 2 Variable loadings on principal components (PCs) and contribution of original variables in discriminant functions (DFs) from external and craniodental measurements, respectively

	PCA		DAPC					
	PC 1	PC 2	PC 1	PC 2	DF 1	DF 2	DF 1	DF 2
HB	**0.843**	0.222	–	–	**0.307**	**0.451**	–	–
T	**0.696**	0.433	–	–	**0.385**	**0.531**	–	–
E	0.255	**0.953**	–	–	0.027	0.004	–	–
HF	**0.749**	0.201	–	–	0.011	0.002	–	–
FA	**0.877**	0.220	–	–	**0.211**	0.007	–	–
Tib	**0.795**	0.242	–	–	0.060	0.003	–	–
GTL	–	–	**0.823**	0.482	–	–	**0.231**	0.001
CCL	–	–	**0.870**	0.447	–	–	**0.210**	0.004
BCW	–	–	0.529	**0.716**	–	–	0.018	0.001
BCH	–	–	0.534	0.386	–	–	0.065	**0.582**
ZYW	–	–	**0.828**	0.459	–	–	0.093	0.013
MAW	–	–	0.364	0.549	–	–	0.008	**0.365**
IOW	–	–	–	**0.925**	–	–	0.002	0.001
CM³L	–	–	**0.868**	0.362	–	–	0.037	0.002
CCW	–	–	**0.922**	0.310	–	–	0.031	0.003
M³M³W	–	–	0.741	0.16	–	–	0.029	0.010
RCM	–	–	**0.834**	−0.103	–	–	0.001	0.001
LC₁M₃L	–	–	**0.859**	0.295	–	–	0.044	0.015
ML	–	–	**0.866**	0.428	–	–	**0.177**	0.001
CPH	–	–	**0.880**	0.358	–	–	0.055	0.004

For abbreviations, see text and Table 1. Bold text indicates high loading of variable on related PC and DF. –: Not available.

power in determining species and sex than the external one, nevertheless both received an ambiguous reclassification in determining *M. shuipuensis*, *M. rongjiangensis*, and our new species and their sex (PP<0.80 for the most likely group and its actual allocation) (Figure 1E–F). These results indicate external and craniodental similarity as well as difficulty in identification using morphological or craniodental measurements alone.

COI alignment spanned 671 bp, including 265 and 256 variable sites and parsimony informative sites, respectively. *Rag2* alignment covered 1 339 bp, including 127 variable sites and 58 parsimony informative sites. For both alignments, the best partitioning scheme selected was one separating each codon position into a partition. For ML analyses, the best nucleotide substitution models for the first, second, and third partitions of *COI* were TIM+G, HKY, and GTR+G, respectively, whereas the best nucleotide substitution models for three position of *Rag2* were TRN+I, HKY, and TIM+G, respectively. The ML trees recovered the genus *Murina* as a well-supported (BS=100) monophyletic group (Figure 1G–H). The phylogenetic reconstructions of the *COI* sequences revealed similar species groups as reported by Francis & Eger (2012) and Eger & Lim (2011). Within the genus, a robust clade, including several *Murina* species known from southern China, emerged (highlighted with gray rectangle in Figure 1G,

BS=100). The new species formed a basal, well-supported, monophyletic group within this clade, which included *M. rongjiangensis*, *M. fanjingshanensis*, *M. shuipuensis, M. hilgendorfi*, and *M. leucogaster*. Due to the limited number of phylogenetically informative variations, phylogeny using *Rag2* resulted in a more ambiguous topology (Figure 1H). However, a well-supported clade (highlighted with gray rectangle in Figure 1H, BS=82), similar to the species assemblage based on the *COI* marker, also emerged, with the basal position of the new species supported (Figure 1GH). Hence, the distinctiveness of the new species was supported by both mitochondrial and nuclear markers.

Taxonomic account

Murina jinchui **sp. nov.** Yu, Csorba, Wu (Figures 1–2; Table 1; Supplementary Table S1)

Common names: Jinchu's Tube-nosed Bat (锦矗管鼻蝠, Jinchu Guanbifu).

Holotype: GZHU 14463, adult male, skin and body in alcohol with skull extracted and cleaned, collected on 13 August 2014 by Yi Wu, Feng Li, Bo-Cheng Chen, and Qiu-Ping Zhang. Deposited in the Key Laboratory of Conservation and Application in Biodiversity of South China, School of Life Sciences, Guangzhou University. The nucleotide sequences of the mitochondrial gene *COI* (GenBank accession No. MN549070) and nuclear gene *Rag2* (GenBank accession No. MN549091) were deposited in GenBank.

Paratypes: All paratypes were collected on 13 August 2014 from the type locality. They were preserved in alcohol with their skulls removed, and deposited in the Key Laboratory of Conservation and Application in Biodiversity of South China, School of Life Sciences, Guangzhou University (GZHU 14453 ♀; GZHU 14454 ♀; GZHU 14455 ♀; GZHU 14462 ♂) and the Hungarian Natural History Museum (GZHU 14461=HNHM 2019.1.1. ♂)

Measurements (in mm) and body mass (in g) of holotype: HB, 40.1; T, 36.0; HF, 7.4; E, 14.4; Tib, 14.7; FA, 32.6; GTL, 15.71; CCL, 13.68; CBL, 15.12; BBW, 7.18; BCH, 7.04; ZYW, 8.49; MAW, 7.55; PL, 7.03; IOW, 4.05; CM³L, 5.32; CCW, 3.88; M³M³W, 5.51; LCM₃L, 5.72; GLM, 10.90; CPH, 3.34; Wt, 4.5.

Type locality: Hetaoping Giant Panda Training Base, Wolong National Nature Reserve, Wenchuan County, Sichuan Province, China (N31°4′23″, E103°13′02″), 1 800 m a.s.l..

Diagnosis: Small species of *Murina* (FA: 32.4–36.4 mm; TIB: 15.7–16.9 mm, Table 1, Figure 1; Supplementary Figure S1 and Table S1). Ears rounded and small without notch on posterior edge (Figure 2B). Plagiopatagium attached near base of toe (Figure 2C). Overall color of dorsum brownish gray with banded appearance (Figure 2D); ventral fur goose gray with two bands (Figure 2E). Skull delicate and braincase not globose (Figure 2F). Both upper incisors visible in lateral view; second upper premolar well-developed compared with corresponding canine and anterior premolar (Figure 2F); mesostyles on first and second upper molars moderately developed, with distinct cusp (Figure 2G).
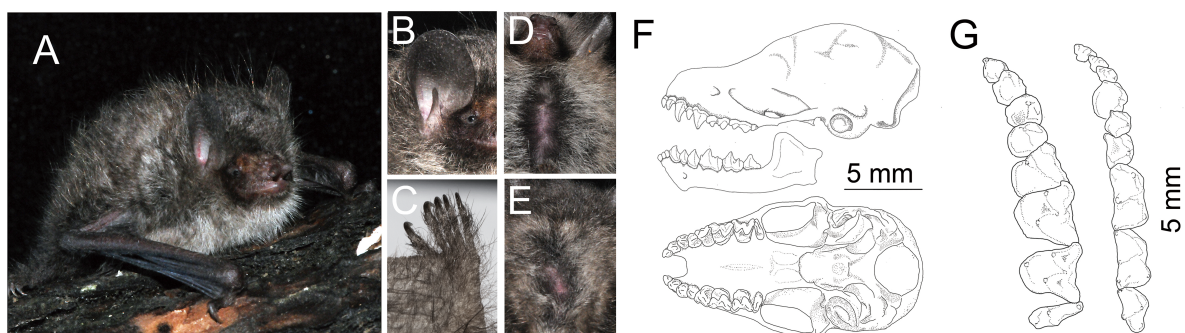
Figure 2 External, skull, and dentition characteristics of *Murina jinchui* sp. nov (holotype, GZHU 14463)
Live individual (A), ear (B), hindfoot (C), dorsal (D) and ventral (E) views of pelage; lateral views of skull and mandible (F); occlusal views of left upper (left) and right lower (right) dentition (G). Scale bars: 5 mm. Photos by Yi Wu (A) and Wen-Hua Yu (B–E).

Description: Small species of *Murina* with '*suilla*-type' dentition. Skin of muzzle, including nose, dark (Figure 2A). Pinnae relatively rounded and large without emargination (Figure 2B). On dorsal surface, basal section of individual hairs black (3 mm, Pantone Process Black C), mid-section brownish gray (1–1.5 mm, Pantone Warm Gray 11C), terminating with dark brown tip (1 mm, Pantone 411C) (Figure 2D). Overall color of dorsum brownish gray (Pantone 410C) (Figure 2A). Underfur overlaid by unicolored warm gray guard hairs (7–8 mm) (3–4 mm, Pantone Warm Gray 11C) (Figure 2D). Upper surface of hind limbs and uropatagium densely furred, particularly along tail and femur, uniformly dark brown. On ventral surface, basal body hairs black (3 mm, Pantone Hexachrome Black C), terminal half progressing to cool gray (Pantone Cool Gray 9 C) (Figure 2E). Ventral guard hairs unicolored gold (7–8 mm, Pantone P7531C) (Figure 2E). Ventral surface of uropatagium covered by short white hairs, rows of papillae with short stiff hairs of similar color as ventral underfur originating from each white dot. Plagiopatagium attached to base of toe. Forearm and metacarpals not furred, short golden hairs present on dorsal surface of thumbs.

Skull small and delicate (Figure 2F; Table 1; Supplementary Table S1), with rostrum sloping gently to forehead (Figure 2F). Laterally, mid-portion of braincase exceeds frontal region in height, braincase not especially domed. No sagittal crest, lambdoid crests relatively weak (Figure 2F). Rostrum deep and pronounced. Depth of nasal emargination exceeds its width, approximately extending to middle of upper canine, outline of emargination varying from smoothly concave to squarish in dorsal view. Basisphenoid pits well-defined, teardrop shaped, elongated, and deep. Zygoma weak, lacking dorsal processes (Figure 2F). Postpalatal emargination with medial projection. Dental formula I 2/3 C 1/1 P 2/2 M 3/3 (Figure 2G). Maxillary toothrows convergent anteriorly (RCM: $\bar{x}$=0.69, range=0.67–0.71, SD=0.02, n=6). I^2 bicuspid, situated anterior to I^3; I^3 with small secondary cusp lingual to primary cusp (Figure 2F). I^2 and I^3 equal in height, first about half of basal area of second upper incisor; second upper incisor in contact with upper canine, approximately half height of C. Basal area of canine equal to that of first premolar, appearing circular in cross-section. P^2 delicate and pointed, about half P^4 in height; P^4 distinctly higher than C, with basal area twice as large. Metacones of M^1 and M^2 exceeding respective paracones in height. Mesostyle of M^1 and M^2 moderately developed but retaining distinct cusp (Figure 2G). Posterior upper molar lacking metacone, with reduced but distinct postparacrista. Mandible delicate, and lower incisors tricuspid. Lower canine exceeding posterior two premolars in height and basal area, well-developed inner cingulum with anterior-most portion in contact with posterior face of I_3. Basal area of P_2 approximately two-thirds that of P_4, height of P_2 approximately equal to that of P_4. First and second lower molars nyctalodont, with well-developed entoconids. Talonids of M_1 and M_2 equal to respective trigonids in crown area, M_3 talonid half of trigonid (Figure 2F, G).

Comparisons with other taxa: *Murina jinchui* **sp. nov.** possesses *suilla*- type dentition having an upper canine smaller basally than the corresponding posterior premolar, and thus is readily distinguished from all species with *cyclotis*-type dental characters. Among all recognized species with *suilla*-type dentition of the phylogenetically most closely related species, *M. hilgendorfi*, *M. leucogaster*, *M. bicolor*, and *M. fanjingshanensis* all are much larger (FA over 37 mm, Table 1); the similar-sized *M. shuipuensis* and *M. rongjiangensis* have an overall reddish dorsal fur, distally yellowish-gold ventral hairs, shorter ears, shorter FA, less convergent maxillary toothrow, and weaker P^4 (Table 1; Supplementary Figure S1 and Table S1).

Other species with a similarly developed P^4, which is basally much larger than C and P^2 (*M. aurata*, *M. harpioloides*, *M. chrysochaetes*, *M. eleryi*, *M. gracilis*, *M. recondita*) have a reddish dorsal pelage, weaker dentition, and domed braincase. With the exception of *M. recondita*, the above species also possess conspicuous shiny golden guard hairs on the dorsum in sharp contrast with the fur of *Murina jinchui* **sp. nov.**

The species *M. feae*, *M. beelzebub*, and *M. jaintiana* are characterized by similar, grayish dorsal and ventral fur, but all have a more domed braincase, and much less developed posterior upper premolars both in height and basal

dimensions.

Etymology: The species is named in honor of Professor Jinchu Hu, one of the founders of the Hetaoping Giant Panda Training Base and Wolong National Nature Reserve in China as well as a well-known giant panda expert, who has supported, encouraged, and actively participated in extensive ecological research in China, especially in Sichuan Province.

Natural history: Currently known only from the type locality, namely Hetaoping Giant Panda Training Base, Wolong National Nature Reserve, Wenchuan County, Sichuan Province, China. The *Murina jinchui* **sp. nov**. specimens were captured in harp traps set in moist evergreen and deciduous broadleaved forests (e.g., *Cyclobalanopsis*, *Quercus*, *Betula*) near the ruins of the original Giant Panda Breeding Center of the Hetaoping Giant Panda Training Base. Three *Rhinolophus ferrumequinum* and one *Pipistrellus* sp. were also caught at this site.

The existence of cryptic diversity within *Murina* is demonstrated by the large number of new species described from the Indomalayan region. China's wide variety of habitats, ranging from tropical to boreal forests and from grassland to desert, has greatly contributed to the richness of chiropteran resources. *Murina* species from China are recorded from temperate regions in the north to subtropical and tropical regions in the south and southwest (Jiang et al., 2015; Liu & Wu, 2019). This vast distribution, together with the many unexplored areas in regard to chiropterological surveys, implies an exceptionally high level of cryptic species diversity in tube-nose bats. Although molecular analysis using DNA barcoding, developed as a tool for rapid identification, can facilitate cryptic species recognition and improve understanding of biogeographical patterns (Csorba et al., 2011; Francis et al., 2010; Francis & Eger, 2012), taxonomic and systematic studies and species identification still depend on morphological data. For instance, species identification of *Murina* typically requires evaluation of morphological characters and availability of properly identified comparative material in museum collections, most of which is deposited in Europe and North America. Such predicaments may be abated in the future by reciprocation of 3D skull models of comparative material using high precision 3D scanners.

Mammalian fur/pelage within many species is subject to a considerable degree of intraspecific color variation (Corbet & Hill, 1992; Davis & Castleberry, 2010; Kries et al., 2018; Son et al., 2015); however, the color patterns of *Murina*, including overall dorsal and ventral aspects, banding of individual hairs, and presence and distribution of brightly colored guard hairs on forearm and/or hindfoot, appear to be stable within species. Therefore, coloration is a useable diagnostic tool for identifying *Murina* species in the field.

Although no sexual or intraspecific variations in pelage color were observed in the new species, sexual size dimorphism, a common phenomenon in *Murina* (Son et al., 2015), emerged during morphological analyses (Figure 1A–C, E; Table 1; Supplementary Figure S1). Females were significantly larger than males in measurements related to length and width of braincase, upper toothrow, and overall size of lower jaw (Table 1). This pattern of sexual dimorphism is suggested to

have occurred under functional limitations and differed among skull regions (e.g., braincase and nasal chambers), therefore doesn't reflect simple allometric size changes.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

W. H. Y., G. C., and Y. W. designed the study; W. H. Y. and Y. W. collected materials for the study; W. H. Y. and Y. W. performed morphometric and phylogenetic analyses; W.H.Y., G.C., and Y.W. interpreted the results and prepared the manuscript, photographs, and figures for the study. All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

Wen-Hua Yu[1], Gabor Csorba[2], Yi Wu[1,*]

[1] *Key Laboratory of Conservation and Application in Biodiversity of South China*, *School of Life Sciences*, *Guangzhou University*, *Guangzhou* 510006, *China*
[2] *Department of Zoology*, *Hungarian Natural History Museum*, *Budapest H*-1088, *Hungary*
*Corresponding author, E-mail: wuyizhouq@263.net

REFERENCES

Chen J, Liu T, Deng HQ, Xiao N, Zhou J. 2017. A new species of *Murina* bats was discovered in Guizhou Province, China. *Cave Research*, **2**(1): 1−10.

Corbet GB, Hill JE. 1992. The Mammals of the Indomalayan Region. London: Natural History Museum.

Csorba G, Son NT, Saveng I, Furey NM. 2011. Revealing cryptic bat diversity: Three new *Murina* and redescription of *M. tubinaris* from Southeast Asia. *Journal of Mammalogy*, **92**(4): 891−904.

Csorba G, Thong VD, Bates PJJ, Furey NM. 2007. Description of a new species of *Murina* from Vietnam (Chiroptera: Vespertilionidae: Murininae). *Occasional Papers, Museum of Texas Tech University*, **268**: 1−9.

Davis AK, Castleberry SB. 2010. Pelage color of red bats *Lasiurus borealis* varies with body size: An image analysis of museum specimens. *Current Zoology*, **56**(4): 401−405.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5): 1792−1797.

Eger JL, Lim BK. 2011. Three new species of *Murina* from southern China (Chiroptera: Vespertilionidae). *Acta Chiropterologica*, **13**(2): 227−243.

Francis CM, Borisenko AV, Ivanova NV, Eger JL, Lim BK, Guillén-Servent

A, Kruskop SV, Mackie I, Hebert PD. 2010. The role of DNA barcodes in understanding and conservation of mammal diversity in southeast Asia. *PLoS One*, **5**: e12575−e12575.

Francis CM, Eger JL. 2012. A review of tube-nosed bats (*Murina*) from Laos with a description of two new species. *Acta Chiropterologica*, **14**(1): 15−38.

Furey NM, Thong VD, Bates PJJ, Csorba G. 2009. Description of a new species belonging to the *Murina* 'suilla-group' (Chiroptera: Vespertilionidae: Murininae) from north Vietnam. *Acta Chiropterologica*, **11**(2): 225−236.

He F, Xiao N, Zhou J. 2015. A new species of *Murina* from China (Chiroptera: Vespertilionidae). *Cave Research*, **2**: 2−6.

Heaney LR, Balete DS, Alviola P, Rickart EA, Ruedi M. 2012. *Nyctalus plancyi* and *Falsistrellus petersi* (Chiroptera: Vespertilionidae) from Northern Luzon, Philippines: Ecology, phylogeny, and biogeographic implications. *Acta Chiropterologica*, **14**(2): 265−278.

Jiang ZG, Ma Y, Wu Y, Wang YX, Zhou KY, Liu SY, Feng Z. 2015. China's Mammal Diversity and Geographic Distribution. Beijing: Science Press. (in Chinese)

Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**(11): 1403−1405.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**: 94.

Kries K, Barros MAS, Duytschaever G, Orkin JD, Janiak MC, Pessoa DMA, Melin AD. 2018. Colour vision variation in leaf-nosed bats (Phyllostomidae): Links to cave roosting and dietary specialization. *Molecular Ecology*, **27**(18): 3627−3640.

Kruskop SV, Eger JL. 2008. A new species of tube-nosed bat *Murina* (Vespertilionidae, Chiroptera) from Vietnam. *Acta Chiropterologica*, **10**(2): 213−220.

Kuo HC, Fang YP, Csorba G, Lee LL. 2009. Three new species of *Murina* (Chiroptera: Vespertilionidae) from Taiwan. *Journal of Mammalogy*, **90**(4): 980−991.

Kuo HC, Soisook P, Ho YY, Csorba G, Wang CN, Rossiter SJ. 2017. A taxonomic revision of the *Kerivoula hardwickii* complex (Chiroptera: Vespertilionidae) with the description of a new species. *Acta Chiropterologica*, **19**(1): 19−39.

Lack JB, Bussche RAVD. 2010. Identifying the confounding factors in resolving phylogenetic relationships in Vespertilionidae. *Journal of Mammalogy*, **91**(6): 1435−1448.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**(6): 1695−1701.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**(3): 772−773.

Liu SY, Wu Y. 2019. Handbook of the Mammals of China. *Fuzhou: Straits Press*: 520.

Maeda K, Matsumura S. 1998. Two new species of Vespertilionid bats, *Myotis* and *Murina* (Vespertilionidae: Chiroptera) from Yanbaru, Okinawa Island, Okinawa Prefecture, Japan. *Zoological Science*, **15**(2): 301−307.

Roehrs ZP, Lack JB, Van Den Bussche RA. 2010. Tribal phylogenetic relationships within Vespertilioninae (Chiroptera: Vespertilionidae) based on mitochondrial and nuclear sequence data. *Journal of Mammalogy*, **91**(5): 1073−1092.

Ruedi M, Biswas J, Csorba G. 2012. Bats from the wet: two new species of Tube-nosed bats (Chiroptera: Vespertilionidae) from Meghalaya, India. *Revue Suisse de Zoologie*, **119**(1): 111−135.

Simmons NB. 2005. Order Chiroptera. *In*: Wilson DE, Reeder DM. Mammal Species of the World: a Taxonomic and Geographic Reference, 3rd edn. Baltimore: The Johns Hopkins University Press, 2142.

Soisook P, Karapan S, Satasook C, Bates PJ. 2013a. A new species of *Murina* (Mammalia: Chiroptera: Vespertilionidae) from peninsular Thailand. *Zootaxa*, **3746**(4): 567−579.

Soisook P, Karapan S, Satasook C, Thong VD, Khan FAA, Maryanto I, Csorba G, Furey N, Aul B, Bates PJJ. 2013b. A review of the *Murina cyclotis* complex (Chiroptera: Vespertilionidae) with descriptions of a new species and subspecies. *Acta Chiropterologica*, **15**(2): 271−292.

Son NT, Csorba G, Tu VT, Thong VD, Wu Y, Harada M, Oshida T, Endo H, Motokawa M. 2015. A new species of the genus *Murina* (Chiroptera: Vespertilionidae) from the central highlands of Vietnam with a review of the subfamily Murininae in Vietnam. *Acta Chiropterologica*, **17**(2): 201−232.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**(5): 758−771.

Zeng X, Chen J, Deng HQ, Xiao N, Zhou J. 2018. A new species of *Murina* from China (Chiroptera: Vespertilionidae). *Ekoloji*, **103**(27): 9−16.

Proximate causes of dispersal for female Yunnan snub-nosed monkeys

DEAR EDITOR,

Individual dispersal trends, unquestionably important for species ecology and evolution, are affected by multiple factors. Understanding the factors that influence female dispersal strategies offers important insight into primate dispersal mechanisms and female choice. To investigate the proximate causes of dispersal in female Yunnan snub-nosed monkeys (*Rhinopithecus bieti*), we observed and analyzed nine years of detailed dispersal and demographic data from a population of *R. bieti* in Xiangguqing, Baimaxueshan Nature Reserve, Yunnan Province, China. Results showed that females who lived long-term in a one-male unit (OMU), without giving birth and with few or no relatives, were more likely to leave that OMU. In addition, an OMU led by an outgroup male and containing more female relatives was significantly more likely to be chosen for immigration. Conversely, greater male age, longer male tenure, and more potentially fertile females discouraged immigration into an OMU. These results suggest that reproduction, male quality, and kin cooperation play the largest roles in female Yunnan snub-nosed monkey dispersal.

Dispersal, defined as the movement of an individual or group between the natal site and another location, is an important factor in the ecology and evolution of many species (Kautz et al., 2016; Nathan, 2001). Over the past four decades, dispersal has been alternately described in terms of mating systems, dispersal patterns, social behaviors, and ecological factors (Dunbar, 1983; Greenwood, 1980; Sterck, 1998; Stokes et al., 2003; Wrangham, 1980). Complex social and ecological dynamics have provided considerable empirical support for explanations of individual dispersal choices in primates and other mammalian taxa (Stokes et al., 2003). Potential factors include infanticide avoidance (Smuts & Smuts, 1993; Stewart & Harcourt, 1987), mate competition (Lawler et al., 2003; Stewart & Harcourt, 1987), inbreeding

prevention (Clutton-Brock, 1989; Packer, 1979; Pusey & Packer, 1987), kin cooperation (Radespiel et al., 2003), local resource competition (Clobert et al., 2001), local resource defense (McNutt, 1996), and dispersal costs (Bonte et al., 2012).

Individual dispersal is ubiquitous among animals, with males typically dispersing more often than females (Greenwood, 1980). Models that account for the benefits to females from dispersal depend on the assumption that females voluntarily move among social groups (Swedell et al., 2011), which they do under certain circumstances (Bowler & Benton, 2005). For example, immigration into a new group can provide access to better resources (Pusey & Packer, 1987; Sterck, 1998). In addition, the competition hypothesis predicts that individuals will disperse if competition is stronger in their current territory than elsewhere (Dobson, 1982). In polygynous primate species, females choose the best available male as a mate due to the high costs of reproduction (Johnstone et al., 1996). Thus, dispersal is often associated with sexual selection and kin cooperation because females can join the group led by their preferred male with lower reproductive investment (Höner et al., 2007).

Long-term studies on the social and ecological mechanisms of dispersal have indicated that the causes of these behaviors in primates and other mammals are multifactorial. As such, it can be difficult to determine the specific reasons for any given dispersal event in non-human primates. However, interaction among these factors undoubtedly influences dispersal. Given the limited information on adult and sub-adult female dispersal events in Yunnan snub-nosed monkeys, we examined the influence of several variables on individual dispersal, including OMU composition, indicators of male quality, and demographic factors. Based on dispersal data collected over nine years of observation, we addressed the following

questions: (1) What factors differ between adult females that emigrated from their OMU and co-resident adult females that did not? (2) What characteristics differentiate females that emigrated from their OMUs from co-resident adult females that did not?

This study was carried out at Xiangguqing in Baimaxueshan Nature Reserve, located on the southern slopes of the Hengduan Mountains (N27°36′, E99°15′), Yunnan Province, China. The study site encompasses an area of almost 90 km^2 and is characterized by a plateau monsoon climate, where temperature and precipitation are strongly seasonal (Li, 2010). Yunnan snub-nosed monkeys are endangered colobines that inhabit high-altitude temperate forests in northwestern Yunnan and southeastern Tibet (Long et al., 1994). This diurnal primate species subsists on lichens and leaves of angiosperms (Li, 2010). They also form large, multilevel social groups consisting of many OMUs and an associated all-male unit (AMU) (Li et al., 2014). Here, the study group consisted of 5–10 OMUs and one AMU each year. Monkeys were individually identifiable using distinctive physical characteristics such as body size, hair pattern, scars, facial features, and pelage color (Ren et al., 2011). As the monkeys were provisioned, they were usually observable through binoculars at distances between 10 m and 20 m. Every day at 0900h and 1700h from 1 January 2010 to 31 December 2018, reserve staff provided food including lichen, carrots, apples, peanuts, and lacquer tree fruit. During those times, we counted the number of individuals in each OMU and AMU (Xia et al., 2016).

In 2008 and 2009, individuals could not yet be reliably identified, so data for those two years were only used to assess age of dispersal for each individual and kinship (mother, daughter, and sister) among individuals within the group. Accurate and detailed records of dispersal, reproduction, and individual mortality were kept from 2010–2018 as part of long-term population monitoring efforts. At every feeding time, in addition to counting the number of individuals in each OMU and AMU, we recorded demographic data including (i) population composition and demographic changes for OMUs and AMU, including births, deaths, and dispersals, and (ii) details of dispersal such as dispersal type, identity of dispersing female, time of dispersal, female age at time of dispersal, age of past and present mates, fertility (female older than four years) before and after dispersal, and number of females relatives in past and present OMUs. Age was assessed by body color, body size, and thinning of white hairs on the back. All dispersal events were confirmed within 1–3 d. The ages of outgroup individuals were estimated using body size and hair color. Females younger than three years and not sexually mature typically migrated with their mothers, and thus were excluded from the scope of this paper. Natal dispersal is primarily considered to be an inbreeding avoidance mechanism and was also excluded from the emigration study.

To explore factors correlated with the likelihood of female dispersal, we divided dispersal into emigration and immigration. Individuals that emigrated were classified as the Experimental Emigration Group (EEG) and individuals in the same OMU that did not emigrate were classified as the Control Emigration Group (CEG). Similarly, the OMUs chosen by immigrant females were identified as the Experimental Immigration Group (EIG) and OMUs not joined by immigrant females were classified as the Control Immigration Group (CIG). We used stepwise binary logistic regression (BLR) to predict female dispersion using reproduction, competition, male/female quality, and kin variables (Table 1). The predictor variables for modelling female emigration included WB (whether female has given birth in the OMU), HLB (length of time female did not breed), FA (female age), FLT (female length of tenure), NR1 (number of relatives), and WOD (whether offspring died within the first year). The predictor variables for modelling female immigration included AM (new alpha male age), MT (new alpha male length of tenure), NFF (number of potentially fertile females in new OMU), NI (total number of individuals in new OMU), OGM (outgroup male became alpha), and NR2 (number of relatives). We included EEG and CEG adult and sub-adult females from the same OMU in the same analyses. We used the Hosmer-Lemeshow goodness-of-fit test and Nagelkerke R^2 measures to determine model fit and significance. We assessed the contribution of predictor variables using the Wald statistic, and the odds ratio (exp(B)) for interpreting the regression models. Significance was set at 0.05 and mean values are presented with standard deviations (±SD). We completed all statistical analyses using SPSS 19.0.

From January 2010 through to December 2018, we recorded a total of 92 female emigration events from 22 OMUs. After excluding dispersal involving individuals younger than three years (19 events), natal dispersal (19 events), and emigration of the whole OMU (13 events), a total of 41 cases of emigration were included in the EEG. Forty-three non-emigration cases occurring in the same OMU at the same time were included in the CEG. The Hosmer-Lemeshow goodness-of-fit test was not significant (X^2=6.21, P=0.624), indicating that the observed data frequencies did not violate the assumptions of the model, and the model was well-fitted (Nagelkerke R^2=0.521). The FLT variable was a significant positive predictor of female emigration (BLR: $β$ ±SE=0.790±0.036, P=0.026), whereas the WB ($β$ ±SE= −2.468±0.923, P=0.007) and NR1 ($β$ ±SE = −1.938±0.537, P<0.001) variables were indicative of a significantly reduced chance of female emigration (Table 2). Using the coefficient values from the final logistic regression output (Table 2), we obtained the logistic regression equation: Y=0.79(FLT)–1.938(NR1)–2.468(WB).

A total of 55 cases of between OMU immigration and three cases of immigration by outgroup individuals were recorded. In eight cases, individuals were younger than three years old and were thus excluded, with the remaining 50 cases used here. The Hosmer-Lemeshow goodness-of-fit test was not significant (X^2=2.489, P=0.962), indicating that the observed data frequencies did not violate the assumptions of the model,

Table 1 Coding and description of variables used in binary logistic regression modelling

Group	Variable type	Variable	Variable description
Emigration Group	Reproduction	WB	Whether female has given birth in OMU (0: no, 1: yes)
		HLB	Length of time female did not breed (after age four)
		WOD	Whether offspring died within first year (0: no, 1: yes)
	Female quality	FLT	Female length of tenure
		FA	Female age
	Kin	NR1	Number of relatives (mother, daughter, and sister)
Immigration Group	Male quality	AM	New alpha male age
		MT	New alpha male length of tenure
		OGM	Outgroup male became alpha (0: no, 1: yes)
	Competition	NFF	Number of potentially fertile females in new OMU
		NI	Total number of individuals in new OMU
	Kin	NR2	Number of relatives (mother, daughter, and sister)

For abbreviations, see text.

Table 2 Binary logistic regression modelling for female emigration and immigration

	Variable type	Variable	Binary logistic regression model						95.0% CI for Exp(B)	
			B	SE	Wald	df	Sig	Exp(B)	Lower	Upper
Emigration	Reproduction	WB	−2.468	0.923	7.150	1	0.007**	0.085	0.014	0.517
		HLB	−0.022	0.055	0.156	1	0.693	0.979	0.879	1.090
		WOD	0.051	1.226	0.002	1	0.967	1.053	0.950	11.638
	Female quality	FA	0.021	0.012	2.993	1	0.084	1.021	0.997	1.046
		FLT	0.79	0.036	4.971	1	0.026*	1.083	1.010	1.161
	Kin	NR1	−1.938	0.537	13.04	1	<0.001**	0.144	0.050	0.412
	Intercept		−0.532	0.881	0.365	1	0.546	0.587	–	–
Immigration	Male quality	AM	−0.036	0.018	3.838	1	0.049*	0.965	1.001	1.079
		MT	−0.106	0.037	8.009	1	0.005**	0.900	0.836	0.968
		OGM	2.048	0.710	8.314	1	0.004**	7.7521	1.927	31.190
	Competition	NFF	−0.991	0.374	7.021	1	0.008**	0.371	0.178	0.773
		NI	0.090	0.181	0.250	1	0.617	1.095	0.768	1.560
	Kin	NR2	4.135	1.016	16.573	1	<0.001**	62.503	8.536	457.662
	Intercept		4.833	2.171	4.954	1	0.026*	125.536	–	–

For abbreviations, see text. B: Logistic coefficient; SE: Standard error of estimate; Wald: Wald chi-square values; df: Degrees of freedom; Sig: Significance; Exp(B): Exponentiated coefficient. CI: Confidence interval. –: Not available. *: $P < 0.05$; **: $P < 0.001$.

and the model was well-fitted (Nagelkerke R^2=0.872). Stepwise analysis of the best-fitting model identified five variables that significantly predicted female immigration (Table 2). In the model, the OGM (BLR: $\beta \pm SE$=2.048±0.710, P=0.004) and NR2 ($\beta \pm SE$=4.135±1.016, $P < 0.001$) variables had a significant positive effect on female immigration, whereas the AM ($\beta \pm SE$=−0.036±0.018, P=0.049), MT ($\beta \pm SE$=−0.106±0.037, P=0.005), and NFF ($\beta \pm SE$=−0.991± 0.374, P=0.008) variables were significant negative predictors of female immigration (Table 2). Using the coefficient values from the final logistic regression output (Table 2), we obtained the logistic regression equation: Y=2.048(OGM)+4.135(NR2) −0.036(AM)−0.106(MT)−0.991(NFF).

Our results showed that R. bieti females who have resided within an OMU for a longer period of time, have not given birth, and have few female relatives within the OMU are more likely to emigrate. Female age, length of time without

breeding, and whether an infant died in its first year had no significant effect on emigration. For factors pertaining to reproduction, only whether a female had given birth determined the likelihood of emigration. Adult females without offspring may be less constrained in their dispersal because there is no risk that the male in a new OMU will kill their existing dependent infant (Smuts & Smuts, 1993; Stewart & Harcourt, 1987). Females who had not given birth for a long time (dispersed and undispersed: 17.1 months vs. 16.5 months) were not more likely to disperse. The average inter-birth interval for R. bieti is two years (Cui et al., 2006), and mating is seasonal, occurring between July to October (Xiang & Sayers, 2009). It is possible that the relatively long reproductive cycle in R. bieti reduces their sensitivity to the length of time since birth. In regard to the influence of infant loss on dispersal, first-year infant mortality during the study period was 15.5% (unpublished data), which may have been

too small to produce significant results. Additionally, as mothers are the primary caregivers of infants, resident males in OMUs may have no direct influence on infant survival apart from infanticide.

Female tenure length and female age were used to explore the relationship between female quality and likelihood of emigration. Only tenure length had a significant effect on female emigration, with a female increasingly likely to emigrate as her tenure in an OMU increased. Longer female tenure in an OMU is correlated with longer tenure of its resident male; for example, longer tenure in western lowland gorillas (*Gorilla gorilla gorilla*) increases the probability of transfer (Stokes et al., 2003). Once they reach an advanced age, the social rank of an alpha male decreases; generally, lower-ranking individuals cannot compete as effectively for resources (Murray et al., 2007; Vogel, 2005). In addition, lower-ranking resident males are more susceptible to displacement by challengers (Zhu et al., 2016). Thus, females may leave an OMU with an older resident male to protect themselves from accidental injury or their infant from infanticide.

Females with more relatives in an OMU were less likely to disperse. In most polygynous non-human primate societies, females exhibit strong kin bonds, often forming matrilineal societies (Pusey & Packer, 1987). The importance of kin in determining offspring dispersal has been demonstrated in multiple species (e.g., *Microtus oeconomus*, Gundersen & Andresassen, 1998; *Microtus townsendii*, Lambin, 1994; *Lacerta vivipara*, Léna et al., 1998). In western gorillas (*Gorilla gorilla*), nearly half of adult females have adult female relatives living in the same group, despite the fact that females disperse out of their natal groups (Bradley et al., 2007). In the absence of female philopatry, continued associations with female kin into adulthood may be the result of choices made during dispersal. Indeed, related female western gorillas live in the same group more frequently than would be expected by chance (Bradley et al., 2007). Our results in *R. bieti*, which exhibit bisexual dispersal similar to other snub-nosed monkeys such as *Rhinopithecus roxellana* (Qi et al., 2009), further support the observation that female kin bonds can be important to social organization, even in primate species without female philopatry and strong matrilines.

Immigration is a selective process, requiring an adult female who has left her OMU to prefer a new, more satisfactory OMU with a high-quality male (Bowler & Benton, 2005). The high cost of reproduction means that in polygynous primates, females are expected to be choosy (Johnstone et al., 1996). In this study, three variables related to male quality (i.e., AM, MT, and OGM) had a significant influence on whether females immigrated into their OMU. Females preferred younger males with a shorter tenure as the OMU leader, as well as outgroup males. Younger resident males had less time to build their OMU, translating to fewer resident females and less female-female competition. Females may prefer mating with outgroup males to enhance genetic diversity (Lehmann & Perrin, 2003). Females also generally avoid mating with familiar males who

were members of their natal group to reduce the chance of costly inbreeding (Höner et al., 2007).

Competition appears to be the primary reason for individual dispersal between social groups. Increasing population density can reduce individual fitness, thus becoming a driving force for dispersal (Bowler & Benton, 2005). Types of competition that emerge as group size increases include kin, resource, and mating competition (Clobert et al., 2001). In this study, the dispersal behavior of *R. bieti* was not consistent with the resource competition hypothesis because the total number of individuals in an OMU (NI) had no significant effect on whether a female immigrated into it. However, the number of fertile females in an OMU (NFF) had a significant negative relationship with female immigration, indicating that females preferred to join OMUs with reduced mating competition, consistent with the mating competition hypothesis. Similar results have been reported for *R. roxellana* (Qi et al., 2009). Yunnan snub-nosed monkey social groups consist of many OMUs and an associated AMU (Li et al., 2014). If the females in an OMU mate exclusively with the single resident male, this creates intense mating competition among them. Additionally, sexual maturity occurs relatively late in *R. bieti* females. Field records indicate that they do not reproduce before the age of five (unpublished data). Considering these facts, in addition to the two-year average inter-birth interval (Cui et al., 2006; Kirkpatrick et al., 1998) and four-month window for conception (Xiang & Sayers, 2009), females in this species must invest considerable effort into reproduction. Therefore, the opportunity for enhanced reproductive success would provide strong motivation for dispersal. Many other primates have been observed to disperse to improve their chances of reproduction (Glander, 1980, 1992;Moore & Ali, 1984; Stewart & Harcourt, 1987; Watts, 1990; Wrangham, 1980). Finally, our results were inconsistent with the predation hypothesis, which predicts that individuals would choose larger OMUs to reduce the risk of predation (Cadet et al., 2003). One reason for this could be the reduced predation pressure due to the decline in monkeys' natural enemies in the families Felidae and Accipitriformes (Li, 2010).

Kin may cooperate to acquire or defend mates or resources, and to prevent unrelated competitors from joining the group (Le Galliard et al., 2003). Choosing an OMU with female relatives allows females to form kin-based alliances, potentially improving their fitness (Dunbar, 1983). Although *R. bieti* females preferentially join OMUs with female relatives, females within OMUs are not strongly genetically related outside of mother, daughter, and sister relationships. This lack of genetic similarity among females is also found in *R. roxellana* (Qi et al., 2009). Kin-based coalitions of female primates can better compete for patchy resources, such as high-quality food and habitat (Wrangham, 1980). Yunnan snub-nosed monkeys inhabit high-altitude temperate forests (Long et al., 1994) and subsist on lichens and mature leaves for much of the year, both of which are of low nutritional value (Li, 2010). In summer and autumn, when higher-quality foods such as fruits and bamboo become available, they are

uniformly distributed (Li, 2010). This helps to explain why this species does not form matrilineal groups to monopolize resources.

In conclusion, our study indicated that multiple factors influenced female dispersal in Yunnan snub-nosed monkeys. These factors were associated with reproduction, infanticide, mate competition, inbreeding avoidance, and kin cooperation but not with local resource defense. The likelihood of female emigration from an OMU increased with her length of tenure and decreased with the number of relatives in the group or if she had given birth. Whether an OMU was led by an outgroup male or contained more female relatives had a significant positive impact on female immigration. In contrast, the chance of female immigration declined with increasing male age, male tenure, and number of fertile females. We argue that female mate choice, inbreeding avoidance, and kin cooperation primarily governed female dispersal in *R. bieti*. In addition, in contrast to other Asian colobines, *R. bieti* does not organize around strong matrilines due to the abundance of low-quality food year-round and uniform distribution of seasonally available high-quality food resources.

SCIENTIFIC FIELD SURVEY PERMISSION INFORMATION

This field experiments were approved by the State Forestry and Grassland Administration of China (the Second National Survey on Terrestrial Wildlife Resources in China).

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

W.C.X., S.N.J., Y.T, and D.Y. L. wrote the manuscript; W.C.X., D.Y.L., X.M. H., and T.Z. collected field data; B.P.R. and A.K. revised the manuscript. All authors read and approved the final version of the manuscript.

Wan-Cai Xia[1,2,3], Sheng-Nan Ji[4], Bao-Ping Ren[5],
Xin-Ming He[6], Tai Zhong[6], Ali Krzton[7], Yun Tang[3,*],
Da-Yong Li[1,*]

[1] *Institute of Rare Animals and Plants, China West Normal University, Nanchong, Sichuan 637009, China*
[2] *School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China*
[3] *Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong, Sichuan 637009, China*
[4] *State Environmental Protection Key Laboratory of Regional Ecological Processes and Functions Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China*
[5] *Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China*
[6] *Baimaxueshan Natural Nature Reserve, Diqing, Yunnan 674400, China*
[7] *RBD Library, Auburn University, AL 36849, USA*
*Corresponding authors, E-mail: tangyun_502@cwnu.edu.cn; 980119lsc@163.com

REFERENCES

Bonte D, Van DH, Bullock JM, Coulon A, Delgado M, Gibbs M, Lehouck V, Matthysen E, Mustin K, Saastamoinen M, Schtickzelle N, Stevens VM, Vandewoestijne S, Baguette M, Barton K, Benton TG, Clobert ACBJ, Dytham C, Hovestadt T, Meier CM, Palmer SCF, Turlure C, Travis JMJ. 2012. Costs of dispersal. *Biological Reviews*, **87**(2): 290−312.

Bowler DE, Benton TG. 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, **80**(2): 205−225.

Bradley BJ, Doransheehy DM, Vigilant L. 2007. Potential for female kin associations in wild western gorillas despite female dispersal. *Proceedings of the Royal Society B Biological Sciences*, **274**(1622): 2179−2185.

Cadet C, Ferrière R, Metz JAJ, van Baalen M. 2003. The evolution of dispersal under demographic stochasticity. *The American Naturalist*, **162**(4): 427−441.

Clobert J, Danchin E, Dhondt AA, Nichols JD. 2001. Migration. Oxford: Oxford University Press.

Clutton-Brock TH. 1989. Female transfer and inbreeding avoidance in social mammals. *Nature*, **337**(6202): 70−72.

Cui LW, Sheng AH, He SC, Xiao W. 2006. Birth seasonality and interbirth interval of captive *Rhinopithecus bieti*. *American Journal of Primatology*, **68**(5): 457−463.

Dobson FS. 1982. Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour*, **30**(4): 1183−1192.

Dunbar RIM. 1983. Structure of gelada baboon reproductive units. II. Social relationships between reproductive females. *Animal Behaviour*, **31**(2): 556−564.

Glander KE. 1980. Reproduction and population growth in free-ranging mantled howling monkeys. *American Journal of Physical Anthropology*, **53**(1): 25−36.

Glander KE. 1992. Dispersal patterns in Costa Rican mantled howling monkeys. *International Journal of Primatology*, **13**(4): 415−436.

Greenwood PJ. 1980. Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**(4): 1140−1162.

Gundersen G, Andreassen HP. 1998. Causes and consequences of natal dispersal in root voles, *Microtus oeconomus*. *Animal Behaviour*, **56**(6): 1355−1366.

Höner OP, Wachter B, East ML, Streich WJ, Wilhelm K, Burke T, Hofer H. 2007. Female mate-choice drives the evolution of male-biased dispersal in a social mammal. *Nature*, **448**(7155): 798−801.

Johnstone RA, Reynolds JD, Deutsch JC. 1996. Mutual mate choice and sex differences in choosiness. *Evolution*, **50**(4): 1382−1391.

Kautz M, Imron MA, Dworschak K, Schopf R. 2016. Dispersal variability and associated population-level consequences in tree-killing bark beetles. *Movement Ecology*, **4**(9): 1−12.

Kirkpatrick RC, Long YC, Zhong T, Xiao L. 1998. Social organization and range use in the Yunnan snub-nosed monkey *Rhinopithecus bieti*.

International Journal of Primatology, **19**(1): 13−51.

Lambin X. 1994. Sex-ratio variation in relation to female philopatry in *Townsend voles*. *Journal of Animal Ecology*, **63**(4): 945−953.

Lawler RR, Richard AF, Riley MA. 2003. Genetic population structure of the white sifaka (*Propithecus verreauxi verreauxi*) at Beza Mahafaly Special Reserve, southwest Madagascar (1992-2001). *Molecular Ecology*, **12**(9): 2307−2317.

Le Galliard JF, Ferrière R, Clobert J. 2003. Mother - offspring interactions affect natal dispersal in a lizard. *Proceedings of the Royal Society B Biological Sciences*, **270**(1520): 1163−1169.

Lehmann L, Perrin N. 2003. Inbreeding avoidance through kin recognition: choosy females boost male dispersal. *The American Naturalist*, **162**(5): 638−652.

Léna JP, Clobert J, De Fraipont M, Lecomte J, Guyot G. 1998. The relative influence of density and kinship on dispersal in the common lizard. *Behavioral Ecology*, **9**(5): 500−507.

Li DY. 2010. Time Budgets, Sleeping Behavior and Diet of the Yunnan Snub- nosed Monkeys (*Rhinopithecus bieti*) at Xiangguqing in Baimaxueshan Nature Reserve (in Chinese). Ph.D. dissertation, Xi'an, China: University of North West University.

Li YH, Li DY, Ren BP, Hu J, Li BG, Krzton A, Li M. 2014. Differences in the activity budgets of Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) by age-sex class at Xiangguqing in Baimaxueshan Nature Reserve, China. *Folia Primatologica*, **85**(6): 335−342.

Long YC, Kirkpatrick CR, Zhong T, Xiao L. 1994. Report on the distribution, population, and ecology of the Yunnan snub-nosed monkey (*Rhinopithecus bieti*). *Primates*, **35**(2): 241−250.

Mcnutt JW. 1996. Sex-biased dispersal in African wild dogs, *Lycaon pictus*. *Animal Behaviour*, **52**(6): 1067−1077.

Moore J, Ali R. 1984. Are dispersal and inbreeding avoidance related?. *Animal Behaviour*, **32**(1): 94−112.

Murray CM, Mane SV, Pusey AE. 2007. Dominance rank influences female space use in wild chimpanzees, *Pan troglodytes*: towards an ideal despotic distribution. *Animal Behaviour*, **74**(6): 1795−1804.

Nathan R. 2001. The challenges of studying dispersal. Trends in Ecology & Evolution. 16(9): 481-483.

Packer C. 1979. Inter-troop transfer and inbreeding avoidance in *Papio anubis*. *Animal Behaviour*, **27**(1): 1−36.

Pusey AE, Packer C. 1987. Migration and philopatry. *In*: Primate Societies. Chicago: University of Chicago Press.

Qi XG, Li BG, Garber PA, Ji W, Watanabe K. 2009. Social dynamics of the golden snub-nosed monkey (*Rhinopithecus roxellana*): female transfer and one-male unit succession. *American Journal of Primatology*, **71**(8): 670−679.

Radespiel U, Lutermann H, Schmelting B, Bruford MW, Zimmermanna E. 2003. Patterns and dynamics of sex-biased dispersal in a nocturnal primate, the grey mouse lemur, *Microcebus murinus*. *Animal Behaviour*, **65**(4): 709−719.

Ren BP, Li DY, He XM, Qiu JH, Li M. 2011. Female resistance to invading males increases infanticide in langurs. *PLoS One*, **6**(4): e18971.

Smuts BB, Smuts RW. 1993. Male aggression and sexual coercion of females in nonhuman primates and other mammals: evidence and theoretical implications. *Advances in the Study of Behavior*, **22**(1): 1−63.

Sterck EHM. 1998. Female dispersal, social organization, and infanticide in langurs: Are they linked to human disturbance?. *American Journal of Primatology*, **44**(4): 235−254.

Stewart KJ, Harcourt AH. 1987. Gorillas: variation in female relationships. *In*: Smuts BB, Cheney DL, Seyfarth RM, Wrangham RW, Struhsaker TT (eds). Primate Societies. Chicago: University of Chicago Press, 155-164.

Stokes EJ, Parnell RJ, Olejniczak C. 2003. Female dispersal and reproductive success in wild western lowland gorillas (*Gorilla gorilla gorilla*). *Behavioral Ecology and Sociobiology*, **54**(4): 329−339.

Swedell L, Saunders J, Schreier A, Davis B, Tesfaye T, Pines M. 2011. Female "dispersal" in hamadryas baboons: transfer among social units in a multilevel society. *American Journal of Physical Anthropology*, **145**(3): 360−370.

Vogel ER. 2005. Rank differences in energy intake rates in white-faced capuchin monkeys (*Cebus capucinus*): the effects of contest competition. *Behavioral Ecology & Sociobiology*, **58**(4): 333−344.

Watts DP. 1990. Ecology of gorillas and its relation to female transfer in mountain gorillas. *International Journal of Primatology*, **11**(1): 21−45.

Wrangham RW. 1980. An ecological model of female-bonded primate groups. *Behaviour*, **75**(3-4): 262−300.

Xia WC, Ren BP, Li YH, Hu J, He XM, Krzton A, Li M, Li DY. 2016. Behavioural responses of Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) to tourists in a provisioned monkey group in Baimaxueshan Nature Reserve. *Folia Primatologica*, **87**(6): 349−360.

Xiang ZF, Sayers K. 2009. Seasonality of mating and birth in wild black-and-white snub-nosed monkeys (*Rhinopithecus bieti*) at Xiaochangdu, Tibet. *Primates*, **50**(1): 50−55.

Zhu PF, Ren BP, Garber PA, Xia F, Grueter CC, Li M. 2016. Aiming low: a resident male's rank predicts takeover success by challenging males in Yunnan snub-nosed monkeys. *American Journal of Primatology*, **78**(9): 974−982.

Identification of a functional 339 bp *Alu* insertion polymorphism in the schizophrenia-associated locus at 10q24.32

DEAR EDITOR,

Genome-wide association studies (GWAS) have identified multiple single nucleotide polymorphisms (SNPs) or small indels robustly associated with schizophrenia; however, the functional risk variations remain largely unknown. We investigated the 10q24.32 locus and discovered a 339 bp *Alu* insertion polymorphism (rs71389983) in complete linkage disequilibrium (LD) with the schizophrenia GWAS risk variant rs7914558. The presence of the *Alu* insertion at rs71389983 strongly repressed transcriptional activities in *in vitro* luciferase assays. This polymorphism may be a target for future mechanistic research. Our study also underlines the importance and necessity of considering previously underestimated *Alu* polymorphisms in future genetic studies of schizophrenia.

Schizophrenia is a severe chronic psychiatric disorder with high heritability (Sullivan et al., 2003), and depicting the genetic architecture of schizophrenia is essential for understanding its pathophysiology. So far, GWAS have identified numerous risk loci (Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), and several studies have attempted to identify causative risk variations and underlying biological mechanisms from the massive tagged single nucleotide polymorphisms (SNPs) (Duan et al., 2014; Huo et al., 2019; Wu et al., 2017, 2019 Yang et al., 2018). However, one potential limitation of current GWAS platforms is that they have primarily focused on SNPs and small indels, ignoring other sequence variations that have also been implicated in the genetic risk of human disorders including schizophrenia (Payer et al., 2017; Song et al., 2018; Yang et al., 2019) and

in non-human primates (Liu et al., 2018). For instance, Song et al. (2018) previously identified a functional human-specific tandem repeat in the *CACNA1C* gene as a potential causative variation for schizophrenia and bipolar disorder.

The chromosomal 10q24.32 region is a critical locus showing genome-wide significant associations with schizophrenia. For example, rs7914558 is reported to be the most significant SNP in the 10q24.32 region in the PGC1 GWAS of European populations (P=1.82×10^{-9}, n=51 695) (Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011), and its association with schizophrenia has been further confirmed in subsequent GWAS with increased sample size (P=3.49×10^{-15}, n=79 845) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Intriguingly, according to data from a recent GWAS of East Asian populations, rs7914558 is also significantly associated with schizophrenia genome-wide (P=3.50×10^{-8}, n=58 140) (Lam et al., 2019). In the present study, through population genetic analyses, *in vitro* luciferase assays, and expression quantitative trait loci (eQTL) data, we identified a functional 339 bp *Alu* insertion polymorphism (rs71389983) within the 9th intron of the *AS3MT* gene in complete LD with rs7914558.

The study protocol was approved by the Institutional Review Board of the Kunming Institute of Zoology (KIZ), Chinese Academy of Sciences (CAS). Informed consent was obtained before any study-related procedures were carried out. Genotyping of rs71389983 and rs7914558 was conducted using polymerase chain reaction (PCR) on 38 European and 39 Han Chinese subjects, with amplicons analyzed using

agarose gel and Sanger sequencing to determine differences in alleles. The PCR primers were: 5'- ATGTAACTGGTATATCC ATCGCCT-3' (forward) and 5'- AGAAGACTCAAACAGATGAAC GGA-3' (reverse) for rs71389983; and 5'-CTCTACTTGCCCC CTTACAGC-3' (forward) and 5'-GAACCGTATCAGTAATCC AACAGA-3' (reverse) for rs7914558.

The HEK293T (human embryonic kidney 293T) and U87MG (human glioblastoma astrocytoma) cell lines used were originally obtained from the Kunming Cell Bank, KIZ, and the Cell Bank of Type Culture Collection of the CAS, respectively. Both cell lines were checked regularly for mycoplasma infection using PCR and microscopy. No cells were found to be contaminated during the study. The HEK293T cells were cultured in a humidified 5% CO_2 incubator at 37 °C in DMEM basic (Dulbecco's Modified Eagle's Medium) (Gibco, USA) supplemented with 10% fetal bovine serum, 1% non-essential amino acids, 1% sodium pyruvate, and 1% penicillin-streptomycin. The U87MG cells were cultured in a humidified 5% CO_2 incubator at 37 °C in MEM (Minimum Essential Medium) supplemented with 10% fetal bovine serum, 1% sodium pyruvate, 2.2 g/L $NaHCO_3$, and 1% penicillin-streptomycin.

For the reporter gene assays, DNA fragments encompassing rs71389983 with either allele were amplified from human genomic DNA using primers 5'-GGCTGCCAGGTTCAAGTAAT- 3' (forward) and 5'-CACACTGGAATACTATTCAGACTT-3' (reverse). The sequences were then cloned into the pGL3-promoter vector (Promega, USA) upstream of the SV40 promoter. The recombinant clones were verified through Sanger sequencing to ensure they only differed at the rs71389983 locus. The pGL3-promoter reporters were transiently co-transfected into cells together with the pRL-TK plasmid (Promega, USA) using Lipofectamine 3000 (Thermo Fisher Scientific, USA). All plasmids were accurately quantified and equal amounts were used for transfection. All transfection procedures lasted 36–48 h, and the cells were then collected to measure luciferase activity using the Dual-Luciferase Reporter Assay System (Promega, USA). The activity of firefly luciferase was normalized to that of Renilla luciferase to control for variations in transfection efficiency. All assays were performed with at least three biological replicates in independent experiments, and statistical analyses were performed by two-tailed t-tests.

We also examined the impacts of risk SNPs on gene mRNA expression using two public RNA-seq brain eQTL datasets, i. e., BrainSeq Phase 2 (http://eqtl.brainseq.org/phase2/eqtl/) and GTEx (https://www.gtexportal.org/) (Collado-Torres et al., 2019; GTEx Consortium et al., 2017). Briefly, from the BrainSeq dataset, we obtained eQTL data of the dorsolateral prefrontal cortex (DLPFC) from 397 individuals, which were calculated using linear regression by covarying diagnosis, gender, genotyping principal components, and expression principal components. From the GTEx dataset, we retrieved the eQTL association results from the frontal cortex (BA9) of 175 subjects, which were calculated using linear regression by covarying genotyping principal components, gender,

genotyping platforms, and additional covariates.

Recent study has shown that a subset of *Alu* insertion polymorphisms exhibit moderate to strong LD (r^2>0.7) with GWAS risk SNPs of complex illnesses (Payer et al., 2017). We therefore examined whether there were *Alu* insertion polymorphisms within the 10q24.32 region. Using public genomic variation databases (i.e., UCSC, http://genome.ucsc. edu/) followed by Sanger sequencing of target regions, we identified an *Alu* insertion polymorphism (339 bp) rs71389983 in intron 9 of *AS3MT* , which was in complete LD with rs7914558 in the Han Chinese and European populations (both r^2=1.00, Figure 1A). The presence of the *Alu* insertion at rs71389983 was linked with the schizophrenia risk G-allele at rs7914558, and therefore may be associated with increased risk of schizophrenia. We note that the frequency of rs71389983 (and rs7914558) showed divergence between the two populations (frequency of *Alu* insertion at rs71389983: 0.423 in Han Chinese vs. 0.605 in Europeans). We also compared the LD structures of the 10q24.32 region between Europeans and East Asians using genotype data from the 1000 Genomes Project (Genomes Project Consortium et al., 2015), and found that the LD structures were relatively similar across distinct populations, despite showing tiny differences (Figure 1A), in agreement with the significant associations of this genomic area in both populations.

The DNA sequence covering rs7914558 is conserved across humans and non-human primates, whereas the *Alu* polymorphism rs71389983 appears to be human-unique. We thus performed bioinformatics functional prediction of rs7914558 using the HaploReg v4.1 dataset (https://pubs.broadinstitute.org/mammals/haploreg/haploreg. php) (Ward & Kellis, 2012). However, we found that it was unlikely located at any DNA segments showing open-chromatin peaks or directly binding to transcription factors or histone markers (e.g., H3K4me1, H3K4me3, H3K9ac, and H3K27ac). On the other hand, *Alu* insertions have been found to affect both transcription and post-transcriptional processes (Häsler & Strub, 2006). Considering that rs71389983 was found in intron 9 of *AS3MT*, we hypothesized that it may be within the enhancer/repressor region of the genome. To test this, we amplified the DNA fragments spanning rs71389983 from individuals carrying different homozygotes (PCR product length: presence of Alu insertion: 589 bp; absence of Alu insertion: 250 bp), and then sub-cloned them into the pGL3 promoter vector. These plasmids were then transfected into the human HEK293T and U87MG cell lines, and reporter gene assays were carried out to examine their regulatory effects. In the HEK293T cells, the transcriptional activity of the pGL3 promoter containing the *Alu* insertion at rs71389983 was significantly lower than that of the promoter without the allele (P<0.000 01, Figure 1B) and that of the empty vector (P<0.000 01). In the U87MG cells, this trend was reproduced and the presence of the *Alu* insertion at rs71389983 corresponded to significantly lower activity of the pGL3 promoter compared with that of the pGL3 promoter without the allele (P<0.000 01, Figure 1B) and the empty vector (P<0.000 01).
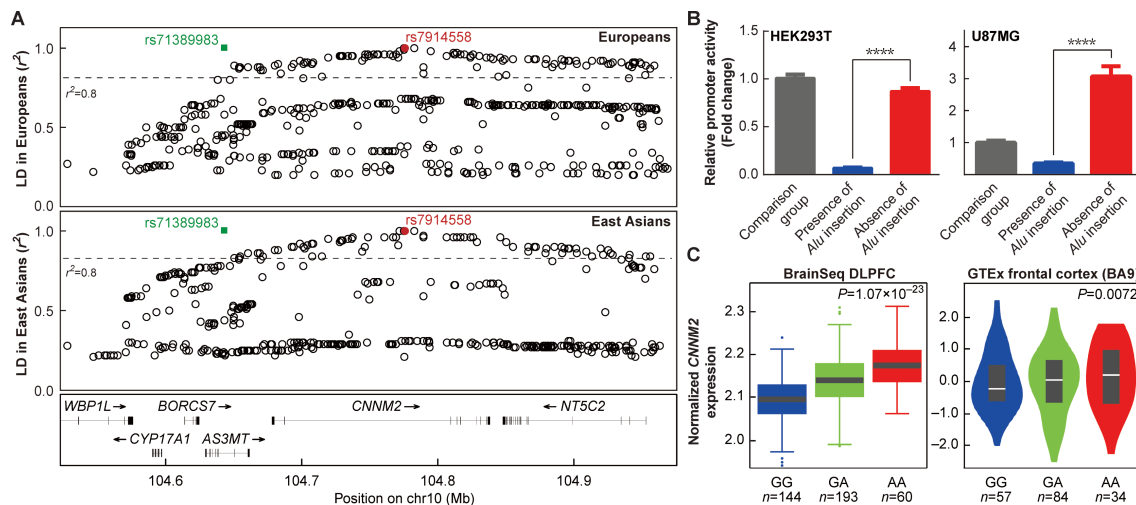
Figure 1 Linkage disequilibrium (LD) analysis of rs7914558 and nearby variations (including rs71389983) in European and East Asian populations (A); Reporter gene assay testing regulatory activity of rs71389983 in HEK293T and U87MG cells (B); Expression quantitative trait loci (eQTL) analyses of rs7914558 with *CNNM2* mRNA in BrainSeq and GTEx datasets (C)

Rs7914558 is located in intron of *CNNM2*, and *Alu* polymorphism rs71389983 is located in intron 9 of *AS3MT*. Both variations are in complete LD in both Europeans and East Asians. Effects of rs71389983 allele variation on pGL3 promoter activity in HEK293T and U87MG cells are shown. "Comparison group" in figure represents empty pGL3 promoter. Values represent fold change in luciferase activity relative to control pGL3 vector. Means and standard deviations of at least three independent experiments are shown. ****: *P*<0.000 01.

Therefore, rs71389983 is likely a functional variation and the *Alu* insertion at this locus likely exerts repressive effects on transcription. In addition to the consistent trend of the effect of the rs71389983 *Alu* insertion on both cell lines, a slight difference between the HEK293T and U87MG cell results was observed, as the pGL3 promoter carrying the "absence of *Alu* insertion" at rs71389983 showed higher transcriptional activity than the empty vector in the U87MG cells, but lower activity than the empty vector in the HEK293T cells. This inconsistency could be explained by the different genetic and physiological backgrounds between the different cell lines.

To further confirm the regulatory effects of the *Alu* polymorphism (rs71389983) on gene expression, we examined two public RNA-seq eQTL datasets (i.e., BrainSeq and GTEx-brain) in human brains (Collado-Torres et al., 2019; GTEx Consortium et al., 2017). As rs71389983 is not genotyped in those eQTL databases, we used rs7914558 as an index SNP. In the BrainSeq dataset, which included DLPFC tissues from 397 individuals, the schizophrenia risk G-allele at rs7914558 was significantly associated with increased gene expression of *BORCS7* (P=9.28×10^{-27}), as well as decreased mRNA expression of *CNNM2* (P=1.07×10^{-23}, Figure 1C) and *CYP17A1-AS1* (P=7.98×10^{-4}). In the 175 frontal cortex (BA9) tissues of the GTEx dataset, the G-allele at rs7914558 was also strongly associated with increased gene expression of *BORCS7* (P=1.00×10^{-9}) and decreased mRNA expression of *CNNM2* (P=0.007 2, Figure 1C), but not with the expression of *CYP17A1-AS1* (P=0.94).

Translating the GWAS risk associations of complex disorders into biological mechanisms remains an urgent task (Barešic et al., 2019; Birnbaum & Weinberger, 2017; Edwards

et al., 2013; Forrest et al., 2018; Gandal et al., 2016). However, most genetic risk loci are located in noncoding regions, which may affect transcription factor binding affinities, gene expression, or even cellular physiological processes (Duan et al., 2014; Forrest et al., 2017; Li et al., 2011; Roussos et al., 2014). We identified a 339 bp *Alu* insertion polymorphism (rs71389983) in the 10q24.32 locus, and reporter gene assays showed that different alleles of rs71389983 exhibited significantly different regulatory activities. The promoter carrying the "absence of *Alu* insertion" at rs71389983 exhibited more than 10-fold higher transcriptional activity than the promoter carrying the "presence of *Alu* insertion", suggesting that the *Alu* insertion sequence likely confers function as a gene silencer. Although this effect is usually caused by certain epigenetic modifications such as DNA methylation or noncoding RNA, the current genome-wide sequencing technologies do not provide ideal tools for answering this question. For example, the ENCODE datasets are mostly based on short DNA reads (≤250 bp) (Encode Project Consortium, 2012), and such methods are not able to precisely map retrotransposons, like *Alu* regions, as *Alu* elements contain multiple highly similar sequences (>300 bp) across the genome. Thus, it is difficult to identify the epigenetic or regulatory markers at rs71389983 (as reflected in the UCSC browser, which shows no ChIP-seq data at the rs71389983 locus). To resolve this problem, long-read sequencing technologies should be applied.

We found that in both the BrainSeq and GTEx-brain tissues, the schizophrenia risk allele at rs71389983 (i.e., its complete linked SNP rs7914558) predicted lower expression of *CNNM2*, consistent with the results of our *in vitro* luciferase assays.

Therefore, *CNNM2* is likely a schizophrenia risk gene, in agreement with previous study (Thyme et al., 2019). However, the present results do not necessarily mean that rs71389983 directly regulates *CNNM2* expression, unless further functional studies (e.g., CRISPR/Cas9 genome editing) are carried out. The significant association of risk SNPs (e.g., rs7914558) at 10q24.32 with *BORCS7* expression is also consistent with earlier research (Duarte et al., 2016; Li et al., 2016a).

Previous studies have demonstrated that *Alu* insertion polymorphisms are significantly associated with multiple complex human disorders and traits, including multiple sclerosis, obesity, height, Alzheimer's disease, breast cancer, and blood pressure (Payer et al., 2017). Our recent study also identified a functional *Alu* polymorphism at 3p21.1 affecting DNA regulatory activity, which was significantly associated with increased risk of psychiatric disorders (e.g., schizophrenia, bipolar disorder, and major depressive disorder) and cognitive disfunctions (Yang et al., 2019). Combined with the present data, these studies suggest that such types of sequence variations may play essential roles in shaping phenotypes during primate or human evolution (Deininger, 2011; Häsler & Strub, 2006). However, our previous study showed that the *Alu* insertion sequence at 3p21.1 increased regulatory activity (Yang et al., 2019), and herein the *Alu* insertion sequence at 10q24.32 reduced transcriptional activities. Although the majority of the *Alu* sequences across the genome show high similarity, their functional regulatory effects may be distinct.

In summary, we discovered a human-unique *Alu* insertion in strong LD with the schizophrenia GWAS risk SNP at 10q24.32. Schizophrenia is hypothesized to be specific to or dominant in humans, and its evolutionary mechanism may be related to unique human variations. For example, we previously identified a human-specific allele rs13107325 at *SLC39A8* undergoing Darwinian natural selection, which enabled humans to adapt to cold environments in Europe, but simultaneously also increased the risk of schizophrenia (Li et al., 2016b). The schizophrenia risk allele at *SLC39A8* is also significantly associated with cognitive function and brain structures in human populations (Davies et al., 2018; Elliott et al., 2018; Luo et al., 2019; Savage et al., 2018). Assuming that the human-unique *Alu* insertions play pivotal roles in shaping humanity, such as development of the dorsolateral prefrontal cortex and higher order human features (e.g., higher cognitive processing) (Wang & Arnsten, 2015), they may also deliver some susceptible or deleterious effects to human health, such as predisposition to schizophrenia. Investigations of such human-unique variations in non-human primates or other species (such as tree shrews) that are evolutionarily close to humans or in human-induced pluripotent stem cells (hiPSC) or reprogrammed cells *via* genome editing, may provide novel insights into the pathophysiology of schizophrenia and other human-dominant disorders (Falk et al., 2016; Hoffman et al., 2019; Luo et al., 2016; Xiao et al., 2017; Xu et al., 2013; Yao, 2017).

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Z.H.Y., Y.L., and M.L. designed the study and interpreted the results. Z.H.Y., N.Q., and X.C. performed DNA extraction, Alu genotyping, population genetic analysis, and primary experiments. L.J.Z., B.L.Z., S.F.Z., J.C., B.X., H.Y.J., D.Y.Z., W.L., H.C., and X.X. contributed to design and helped in the experiments. Z.H.Y., X.X., and M.L. drafted the manuscript.

Zhi-Hui Yang[1,2,#], Xin Cai[1,2,#], Na Qu[3,4,#], Li-Juan Zhao[1,2], Bao-Liang Zhong[3,4], Shu-Fang Zhang[3,4], Jing Chen[3,4], Bin Xia[3,4], Hong-Yan Jiang[5], Dan-Yang Zhou[1,2], Wei-Peng Liu[1,2], Hong Chang[1], Xiao Xiao[1], Yi Li[3,4,*], Ming Li[1,2,6,*]

[1] *Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan 650223*, *China*
[2] *Kunming College of Life Science*, *University of Chinese Academy of Sciences*, *Kunming*, *Yunnan 650223*, *China*
[3] *Affiliated Wuhan Mental Health Center*, *Tongji Medical College*, *Huazhong University of Science and Technology*, *Wuhan*, *Hubei 430012*, *China*
[4] *Research Center for Psychological and Health Sciences*, *China University of Geosciences*, *Wuhan*, *Hubei 430012*, *China*
[5] *Department of Psychiatry*, *First Affiliated Hospital of Kunming Medical University*, *Kunming*, *Yunnan 650032*, *China*
[6] *Center for Excellence in Brain Science and Intelligence Technology*, *Chinese Academy of Sciences*, *Shanghai 200031*, *China*
[#]Authors contributed equally to this work
*Corresponding authors, E-mail: psylee@163.com; limingkiz@mail.kiz.ac.cn

REFERENCES

Barešić A, Nash AJ, Dahoun T, Howes O, Lenhard B. 2020. Understanding the genetics of neuropsychiatric disorders: the potential role of genomic regulatory blocks. *Molecular Psychiatry*, **25**: 6−18.

Birnbaum R, Weinberger DR. 2017. Genetic insights into the neurodevelopmental origins of schizophrenia. *Nature Reviews Neuroscience*, **18**(12): 727−740.

Collado-Torres L, Burke EE, Peterson A, Shin J, Straub RE, Rajpurohit A, Semick SA, Ulrich WS, BrainSeq C, Price AJ, Valencia C, Tao R, Deep-Soboslay A, Hyde TM, Kleinman JE, Weinberger DR, Jaffe AE. 2019. Regional heterogeneity in gene expression, regulation, and coherence in the frontal cortex and hippocampus across development and schizophrenia. *Neuron*, **103**(2): 203−216.

Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, Hagenaars SP, Ritchie SJ, Marioni RE, Fawns-Ritchie C, Liewald DCM, Okely JA, Ahola-Olli AV, Barnes CLK, Bertram L, Bis JC, Burdick KE, Christoforou A, DeRosse P, Djurovic S, Espeseth T, Giakoumaki S, Giddaluru S, Gustavson DE, Hayward C, Hofer E, Ikram MA, Karlsson R,

Knowles E, Lahti J, Leber M, Li S, Mather KA, Melle I, Morris D, Oldmeadow C, Palviainen T, Payton A, Pazoki R, Petrovic K, Reynolds CA, Sargurupremraj M, Scholz M, Smith JA, Smith AV, Terzikhan N, Thalamuthu A, Trompet S, van der Lee SJ, Ware EB, Windham BG, Wright MJ, Yang J, Yu J, Ames D, Amin N, Amouyel P, Andreassen OA, Armstrong NJ, Assareh AA, Attia JR, Attix D, Avramopoulos D, Bennett DA, Böhmer AC, Boyle PA, Brodaty H, Campbell H, Cannon TD, Cirulli ET, Congdon E, Conley ED, Corley J, Cox SR, Dale AM, Dehghan A, Dick D, Dickinson D, Eriksson JG, Evangelou E, Faul JD, Ford I, Freimer NA, Gao H, Giegling I, Gillespie NA, Gordon SD, Gottesman RF, Griswold ME, Gudnason V, Harris TB, Hartmann AM, Hatzimanolis A, Heiss G, Holliday EG, Joshi PK, Kähönen M, Kardia SLR, Karlsson I, Kleineidam L, Knopman DS, Kochan NA, Konte B, Kwok JB, Le Hellard S, Lee T, Lehtimäki T, Li SC, Liu T, Koini M, London E, Longstreth WT Lopez, Loukola A, Luck T, Lundervold AJ, Lundquist A, Lyytikäinen LP, Martin NG, Montgomery GW, Murray AD, Need AC, Noordam R, Nyberg L, Ollier W, Papenberg G, Pattie A, Polasek O, Poldrack RA, Psaty BM, Reppermund S, Riedel-Heller SG, Rose RJ, Rotter JI, Roussos P, Rovio SP, Saba Y, Sabb FW, Sachdev PS, Satizabal CL, Schmid M, Scott RJ, Scult MA, Simino J, Slagboom PE, Smyrnis N, Soumaré A, Stefanis NC, Stott DJ, Straub RE, Sundet K, Taylor AM, Taylor KD, Tzoulaki I, Tzourio C, Uitterlinden A, Vitart V, Voineskos AN, Kaprio J, Wagner M, Wagner H, Weinhold L, Wen KH, Widen E, Yang Q, Zhao W, Adams HHH, Arking DE, Bilder RM, Bitsios P, Boerwinkle E, Chiba-Falek O, Corvin A, De Jager PL, Debette S, Donohoe G, Elliott P, Fitzpatrick AL, Gill M, Glahn DC, Hägg S, Hansell NK, Hariri AR, Ikram MK, Jukema JW, Vuoksimaa E, Keller MC, Kremen WS, Launer L, Lindenberger U, Palotie A, Pedersen NL, Pendleton N, Porteous DJ, Räikkönen K, Raitakari OT, Ramirez A, Reinvang I, Rudan I, Dan R, Schmidt R, Schmidt H, Schofield PW, Schofield PR, Starr JM, Steen VM, Trollor JN, Turner ST, Van Duijn CM, Villringer A, Weinberger DR, Weir DR, Wilson JF, Malhotra A, McIntosh AM, Gale CR, Seshadri S, Mosley TH Bressler, Lencz T, Deary IJ. 2018. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, 9(1): 2098.

Deininger P. 2011. Alu elements: know the SINEs. *Genome Biology*, 12(12): 236.

Duan J, Shi J, Fiorentino A, Leites C, Chen X, Moy W, Chen J, Alexandrov BS, Usheva A, He D, Freda J, O'Brien NL, Molecular Genetics of Schizophrenia Collaboration, Genomic Psychiatric Cohort Consortium, McQuillin A, Sanders AR, Gershon ES, DeLisi LE, Bishop AR, Gurling HM, Pato MT, Levinson DF, Kendler KS, Pato CN, Gejman PV. 2014. A rare functional noncoding variant at the GWAS-implicated MIR137/MIR2682 locus might confer risk to schizophrenia and bipolar disorder. *American Journal of Human Genetics*, 95(6): 744−753.

Duarte RRR, Troakes C, Nolan M, Srivastava DP, Murray RM, Bray NJ. 2016. Genome-wide significant schizophrenia risk variation on chromosome 10q24 is associated with altered cis-regulation of BORCS7, AS3MT, and NT5C2 in the human brain. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics*, 171(6): 806−814.

Edwards SL, Beesley J, French JD, Dunning AM. 2013. Beyond GWASs: illuminating the dark road from association to function. *American Journal of Human Genetics*, 93(5): 779−797.

Elliott LT, Sharp K, Alfaro-Almagro F, Shi S, Miller KL, Douaud G, Marchini J, Smith SM. 2018. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature*, 562(7726): 210−216.

Encode Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414): 57−74.

Falk A, Heine VM, Harwood AJ, Sullivan PF, Peitz M, Brüstle O, Shen S, Sun YM, Glover JC, Posthuma D, Djurovic S. 2016. Modeling psychiatric disorders: from genomic findings to cellular phenotypes. *Molecular Psychiatry*, 21(9): 1167−1179.

Forrest MP, Parnell E, Penzes P. 2018. Dendritic structural plasticity and neuropsychiatric disease. *Nature Reviews Neuroscience*, 19(4): 215−234.

Forrest MP, Zhang H, Moy W, McGowan H, Leites C, Dionisio LE, Xu Z, Shi J, Sanders AR, Greenleaf WJ, Cowan CA, Pang ZP, Gejman PV, Penzes P, Duan J. 2017. Open chromatin profiling in hiPSC-derived neurons prioritizes functional noncoding psychiatric risk variants and highlights neurodevelopmental loci. *Cell Stem Cell*, 21(3): 305−318. e8.

Gandal MJ, Leppa V, Won H, Parikshak NN, Geschwind DH. 2016. The road to precision psychiatry: translating genetics into disease mechanisms. *Nature Neuroscience*, 19(11): 1397−1407.

Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. 2015. A global reference for human genetic variation. *Nature*, 526(7571): 68−74.

GTEx Consortium, Laboratory Data Analysis, Coordinating Center-Analysis Working Group, Statistical Methods Groups-Analysis Working Group, Enhancing GTEx Groups, NIH Common Fund, Nih/Nci, Nih/Nhgri, Nih/Nimh, Nih/Nida, Biospecimen Collection Source Site Ndri, Biospecimen Collection Source Site Rpci, Biospecimen Core Resource Vari, Brain Bank Repository-University of Miami Brain Endowment Bank, Leidos Biomedical-Project Management, Elsi Study, Genome Browser Data Integration, Visualization EBI, Genome Browser Data Integration, Visualization-Ucsc Genomics Institute, University of California Santa Cruz, Lead analysts, Laboratory Data Analysis &Coordinating Center, NIH program management, Biospecimen collection, Pathology: eQTL manuscript working group, Battle A, Brown CD, Engelhardt BE, Montgomery SB. 2017. Genetic effects on gene expression across human tissues. *Nature*, 550(7675): 204−213.

Häsler J, Strub K. 2006. Alu elements as regulators of gene expression. *Nucleic Acids Research*, 34(19): 5491−1597.

Hoffman GE, Schrode N, Flaherty E, Brennand KJ. 2019. New considerations for hiPSC-based models of neuropsychiatric disorders. *Molecular Psychiatry*, 24(1): 49−66.

Huo Y, Li S, Liu J, Li X, Luo XJ. 2019. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nature Communications*, 10(1): 670.

Lam M, Chen CY, Li Z, Martin AR, Bryois J, Ma X, Gaspar H, Ikeda M, Benyamin B, Brown BC, Liu R, Zhou W, Guan L, Kamatani Y, Kim SW, Kubo M, Kusumawardhani A, Liu CM, Ma H, Periyasamy S, Takahashi A, Xu Z, Yu H, Zhu F, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Indonesia Schizophrenia Consortium, Genetic REsearch on schizophreniA neTwork-China and the Netherlands (GREAT-CN), Chen WJ, Faraone S, Glatt SJ, He L, Hyman SE, Hwu HG, McCarroll SA, Neale BM, Sklar P, Wildenauer DB, Yu X, Zhang D, Mowry BJ, Lee J, Holmans P, Xu S, Sullivan PF, Ripke S, O'Donovan MC, Daly MJ, Qin S, Sham P, Iwata N, Hong KS, Schwab SG, Yue W, Tsuang M, Liu J, Ma X, Kahn RS, Shi Y, Huang H. 2019. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics*, 51: 1670−1678.

Li M, Jaffe AE, Straub RE, Tao R, Shin JH, Wang Y, Chen Q, Li C, Jia Y, Ohi K, Maher BJ, Brandon NJ, Cross A, Chenoweth JG, Hoeppner DJ, Wei H, Hyde TM, McKay R, Kleinman JE, Weinberger DR. 2016a. A human-specific AS3MT isoform and BORCS7 are molecular risk factors in the

10q24.32 schizophrenia-associated locus. *Nature Medicine*, **22**(6): 649−656.

Li M, Luo XJ, Xiao X, Shi L, Liu XY, Yin LD, Diao HB, Su B. 2011. Allelic differences between Han Chinese and Europeans for functional variants in ZNF804A and their association with schizophrenia. *The American Journal of Psychiatry*, **168**(12): 1318−1325.

Li M, Wu DD, Yao YG, Huo YX, Liu JW, Su B, Chasman DI, Chu AY, Huang T, Qi L, Zheng Y, CHARGE Nutrition Working Group, DietGen Consortium, Luo XJ. 2016b. Recent positive selection drives the expansion of a schizophrenia risk nonsynonymous variant at SLC39A8 in Europeans. *Schizophrenia Bulletin*, **42**(1): 178−190.

Liu SX, Hou W, Zhang XY, Peng CJ, Yue BS, Fan ZX, Li J. 2018. Identification and characterization of short tandem repeats in the Tibetan macaque genome based on resequencing data. *Zoological Research*, **39**(4): 291−300.

Luo Q, Chen Q, Wang W, Desrivieres S, Quinlan EB, Jia T, Macare C, Robert GH, Cui J, Guedj M, Palaniyappan L, Kherif F, Banaschewski T, Bokde ALW, Buchel C, Flor H, Frouin V, Garavan H, Gowland P, Heinz A, Ittermann B, Martinot JL, Artiges E, Paillere-Martinot ML, Nees F, Orfanos DP, Poustka L, Frohner JH, Smolka MN, Walter H, Whelan R, Callicott JH, Mattay VS, Pausova Z, Dartigues JF, Tzourio C, Crivello F, Berman KF, Li F, Paus T, Weinberger DR, Murray RM, Schumann G, Feng J. 2019. Association of a schizophrenia-risk nonsynonymous variant with putamen volume in adolescents: a voxelwise and genome-wide association study. *JAMA Psychiatry*, **76**(4): 435−445.

Luo X, Li M, Su B. 2016. Application of the genome editing tool CRISPR/Cas9 in non-human primates. *Zoological Research*, **37**(4): 214−219.

Payer LM, Steranka JP, Yang WR, Kryatova M, Medabalimi S, Ardeljan D, Liu C, Boeke JD, Avramopoulos D, Burns KH. 2017. Structural variants caused by Alu insertions are associated with risks for many human diseases. *Proceedings of the Natlional Academy of Sciences of the United States of America*, **114**(20): E3984−E3992.

Roussos P, Mitchell AC, Voloudakis G, Fullard JF, Pothula VM, Tsang J, Stahl EA, Georgakopoulos A, Ruderfer DM, Charney A, Okada Y, Siminovitch KA, Worthington J, Padyukov L, Klareskog L, Gregersen PK, Plenge RM, Raychaudhuri S, Fromer M, Purcell SM, Brennand KJ, Robakis NK, Schadt EE, Akbarian S, Sklar P. 2014. A role for noncoding variation in schizophrenia. *Cell Reports*, **9**(4): 1417−1429.

Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, Nagel M, Awasthi S, Barr PB, Coleman JRI, Grasby KL, Hammerschlag AR, Kaminski JA, Karlsson R, Krapohl E, Lam M, Nygaard M, Reynolds CA, Trampush JW, Young H, Zabaneh D, Hägg S, Hansell NK, Karlsson IK, Linnarsson S, Montgomery GW, Muñoz-Manchado AB, Quinlan EB, Schumann G, Skene NG, Webb BT, White T, Arking DE, Avramopoulos D, Bilder RM, Bitsios P, Burdick KE, Cannon TD, Chiba-Falek O, Christoforou A, Cirulli ET, Congdon E, Corvin A, Davies G, Deary IJ, DeRosse P, Dickinson D, Djurovic S, Donohoe G, Conley ED, Eriksson JG, Espeseth T, Freimer NA, Giakoumaki S, Giegling I, Gill M, Glahn DC, Hariri AR, Hatzimanolis A, Keller MC, Knowles E, Koltai D, Konte B, Lahti J, Le Hellard S, Lencz T, Liewald DC, London E, Lundervold AJ, Malhotra AK, Melle I, Morris D, Need AC, Ollier W, Palotie A, Payton A, Pendleton N, Poldrack RA, Räikkönen K, Reinvang I, Roussos P, Rujescu D, Sabb FW,

Scult MA, Smeland OB, Smyrnis N, Starr JM, Steen VM, Stefanis NC, Straub RE, Sundet K, Tiemeier H, Voineskos AN, Weinberger DR, Widen E, Yu J, Abecasis G, Andreassen OA, Breen G, Christiansen L, Debrabant B, Dick DM, Heinz A, Hjerling-Leffler J, Ikram MA, Kendler KS, Martin NG, Medland SE, Pedersen NL, Plomin R, Polderman TJC, Ripke S, van der Sluis S, Sullivan PF, Vrieze SI, Wright MJ, Posthuma D. 2018. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, **50**(7): 912−919.

Schizophrenia Psychiatric Genome-Wide Association Study Consortium. 2011. Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*, **43**(10): 969−976.

Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, **511**(7510): 421−427.

Song JHT, Lowe CB, Kingsley DM. 2018. Characterization of a human-specific tandem repeat associated with Bipolar Disorder and Schizophrenia. *American Journal of Human Genetics*, **103**(3): 421−430.

Sullivan PF, Kendler KS, Neale MC. 2003. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, **60**(12): 1187−1192.

Thyme SB, Pieper LM, Li EH, Pandey S, Wang Y, Morris NS, Sha C, Choi JW, Herrera KJ, Soucy ER, Zimmerman S, Randlett O, Greenwood J, McCarroll SA, Schier A. F. 2019. Phenotypic landscape of schizophrenia-associated genes defines candidates and their shared functions. *Cell*, **177**(2): 478−491. e20.

Wang M, Arnsten AF. 2015. Physiological approaches to understanding molecular actions on dorsolateral prefrontal cortical neurons underlying higher cognitive processing. *Zoological Research*, **36**(6): 314−319.

Ward LD, Kellis M. 2012. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research*, **40** (Database issue): D930−D934.

Wu Y, Bi R, Zeng C, Ma C, Sun C, Li J, Xiao X, Li M, Zhang DF, Zheng P, Sheng N, Luo XJ, Yao YG. 2019. Identification of the primate-specific gene BTN3A2 as an additional schizophrenia risk gene in the MHC loci. *EbioMedicine*, **44**: 530−541.

W u, Y, Yao YG, Luo XJ. 2017. SZDB: A database for schizophrenia genetic research. *Schizophrenia Bulletin*, **43**(2): 459−471.

Xiao J, Liu R, Chen CS. 2017. Tree shrew (*Tupaia belangeri*) as a novel laboratory disease animal model. *Zoological Research*, **38**(3): 127−137.

Xu L, Zhang Y, Liang B, Lu LB, Chen CS, Chen YB, Zhou JM, Yao YG. 2013. Tree shrews under the spot light: emerging model of human diseases. *Zoological Research*, **34**(2): 59−69.

Yang CP, Li X, Wu Y, Shen Q, Zeng Y, Xiong Q, Wei M, Chen C, Liu J, Huo Y, Li K, Xue G, Yao YG, Zhang C, Li M, Chen Y, Luo XJ. 2018. Comprehensive integrative analyses identify GLT8D1 and CSNK2B as schizophrenia risk genes. *Nature Communications*, **9**(1): 838.

Yang Z, Zhou D, Li H, Cai X, Liu W, Wang L, Chang H, Li M, Xiao X. 2020. The genome-wide risk alleles for psychiatric disorders at 3p21.1 show convergent effects on mRNA expression, cognitive function, and mushroom dendritic spine. *Molecular Psychiatry*, **25**: 48−66.

Yao YG. 2017. Creating animal models, why not use the Chinese tree shrew (*Tupaia belangeri chinensis*)?. *Zoological Research*, **38**(3): 118−126.

Long-term protection against dengue viruses in mice conferred by a tetravalent DNA vaccine candidate

DEAR EDITOR,

The development of an effective tetravalent vaccine against dengue viruses (DENVs) has become a world priority. We previously showed that four monovalent dengue DNA vaccines expressing premembrane (prM) and envelope (E) proteins displayed effective protection against corresponding challenges in mice. Thus, to elucidate the overall immunity and persistence of the tetravalent formulation (TetraME), we evaluated the humoral and cellular immune responses as well as the long-term protection in the current study. TetraME-immunized mice displayed increased production of Th1/Th2-typed cytokines upon stimulation with heterologous DENV antigens. Moreover, high levels of tetravalent DENV antibodies and sterilized immunity were detected long-term (30 weeks after immunization). These findings provide feasible validation for the potential utility of this vaccine formulation.

DENVs are mosquito-borne flaviviruses. The DENV genome contains positive single-stranded RNA encoding three structure proteins and seven non-structure proteins. Among them, prM and envelope E proteins contain epitopes of both cellular immunity and neutralizing antibodies and are therefore often used as molecular targets for vaccine development (Wang et al., 2018). DENVs have four distinct serotypes (DENV1–4) and infection by any serotype can cause dengue fever and/or life-threatening dengue diseases. Recently, due to mosquito-favorable factors such as global warming and increased human population movement, the incidence of dengue is on the rise worldwide (Wilder-Smith et al., 2019) and has become a global public health concern.

Vaccination is the most effective approach against dengue and has been the focus of virologists for many years. Common dengue candidate vaccines are predominantly composed of either live-attenuated or recombinant chimeric vaccines (Shrivastava et al., 2017), with many still under development. At present, Dengvaxia is the only licensed vaccine against all four serotypes of DENV. However, the World Health Organization (WHO) has recommended it only be used in populations previously exposed to DENVs, indicating its limitation in application (Lee et al., 2018). An ideal tetravalent dengue vaccine avoids interference among components and provides long-term and balanced protection against all four serotypes (Fatima & Syed, 2018; Prompetchara et al., 2019). DNA vaccines offer a series of advantages, such as mobilizing the cellular and humoral arms of the immune response and providing prolonged protection against a range of pathogens (Fynan et al., 2018). We previously manufactured four constructs expressing each DENV prM and E proteins, named pV-D1ME–pV-D4ME, which were individually evaluated in regard to immunogenicity and protection in BALB/c mice (Chen et al., 2016; Sheng et al., 2019; Zheng et al., 2017). In the current study, immunocompetent BALB/c mice were vaccinated with four monovalent prM/E-based DNA vaccine candidates (TetraME), after which we investigated the balanced and long-term tetravalent protection. All animal experiments were performed under approval of the Animal Experiments and Experimental Animal Welfare Committee of Chinese Capital Medical University (AEEI-2015-066). All animal experiments were performed under diethyl ether anesthesia, and all efforts were made to minimize suffering.

As shown in Figure 1A, mice were thrice immunized with 50 µg of each monovalent vaccine or pV vector into the quadricep muscles of all four limbs via electroporation (EP) at three-week intervals. To characterize the production of Th2 (IL-4)/Th1 (IFN-γ)-type cytokines in response to DENV1–4, splenocytes harvested from mice one week after final immunization with either TetraME or pV were plated at 3×10^5 cells per well in pre-coated enzyme-linked immunospot plates to quantitatively measure IL-4 or IFN-γ expression. When individually stimulated with DENV1–4 antigens, significantly secreted and comparable levels of the two cytokines were observed in the TetraME vaccination groups compared with

the control groups (Figure 1B, *P*<0.05 or *P*<0.01). The higher levels of IL-4 and IFN-γ indicated functional cytotoxic T cell activity, which contributed to the clearance of virus-infected

cells. These cytokine results suggest that vaccination with TetraME elicited Th1/Th2 mixed immune responses to DENV1–4.
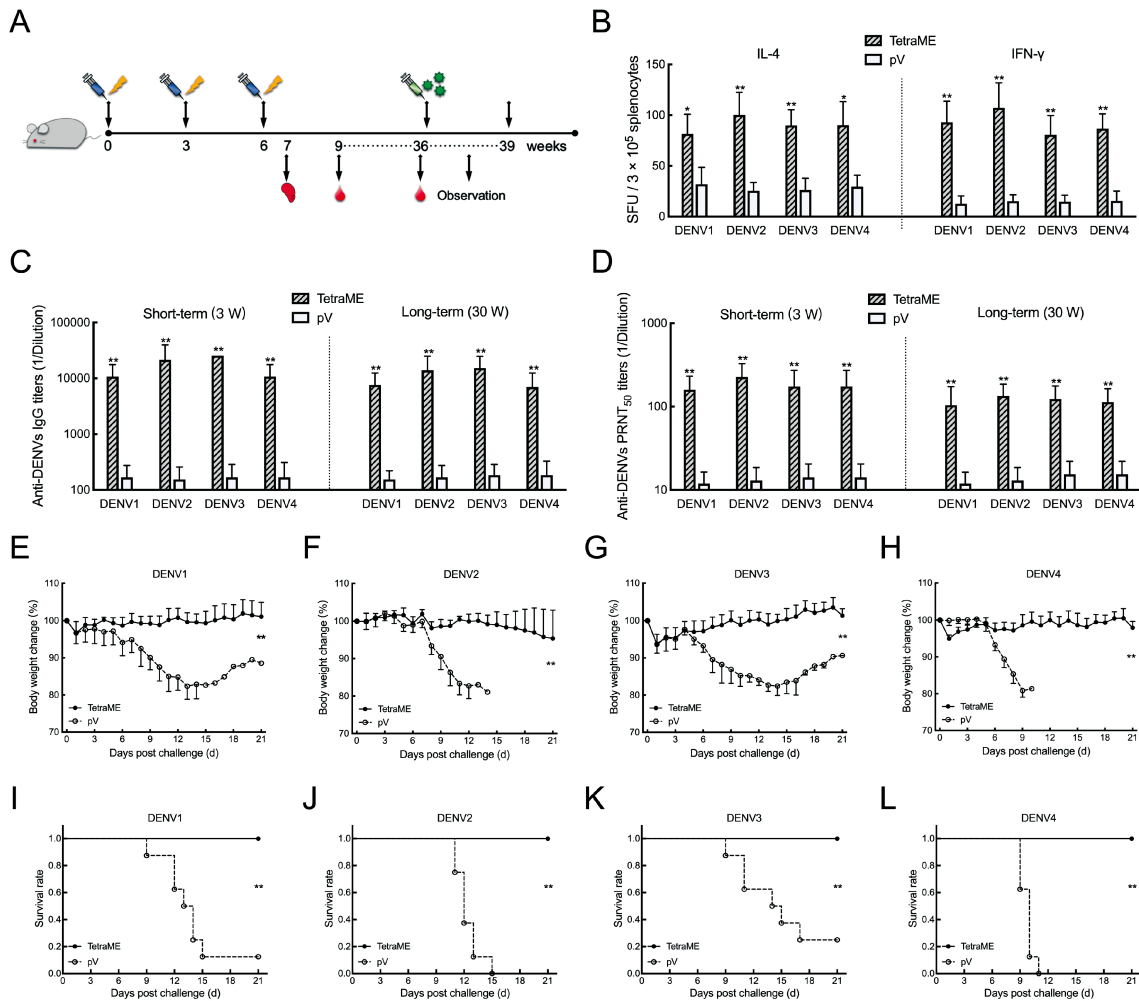

Figure 1 Tetravalent dengue DNA vaccine candidate (TetraME) induces cytokine and short- and long-term humoral immune responses and provides protection against four serotypes of DENV in BALB/c mice

A: Mouse experimental workflow. Groups of mice were immunized by intramuscular electroporation with 50 µg of either monovalent dengue DNA vaccine candidate or vector (pV) in each limb individually and were boosted twice at three-week intervals. Splenocytes were obtained one week after final immunization, and sera were collected three and 30 weeks after final immunization, respectively. Subsequently, vaccinated mice were challenged with 1×10^6 PFU of DENV1, 200 PFU of DENV2, 1×10^6 PFU of DENV3, or 1×10^5 PFU of DENV4. Body weight changes and survival rates were observed for 21 consecutive days after challenge. B: Splenocyte-secreted IL-4 and IFN-γ upon DENV1-4 antigen stimulation. SFU: Spot forming unit. Short- and long-term IgG antibodies (C) and nAb titers (D) against DENV1-4 in sera of immunized mice. Body weight changes (E–H) and survival rates (I–L) were monitored daily for 21 consecutive days. Mice exhibiting more than 20% loss in weight were humanely euthanized for ethical reasons. *n*=8, results are expressed as means±*SD*. *: *P*<0.05; **: *P*<0.01.

To detect short- and long-term DENV-specific IgG antibodies, enzyme-linked immunosorbent assays was used three- and 30-weeks post-vaccination, as described previously (Wang et al., 2018). The TetraME vaccination generated high levels of tetravalent IgG antibodies three weeks after the last immunization. The IgG antibody titers towards DENV1, DENV2, DENV3, and DENV4 were 1:10 763, 1: 21 527, 1: 25

600, and 1: 10 763, respectively, and were significantly different from their corresponding controls (Figure 1C, *P*<0.01). Notably, the IgG antibodies in the sera of immunized mice remained elevated with titers of 1: 7 610, 1: 13 959, 1:15 222, and 1:6 979, respectively, at 30-weeks post-vaccination, indicating a decreasing trend with time but relatively long-term immunogenicity.

Generally, the level of DENV-specific neutralizing antibodies (nAbs) can be used as a predictor of protective immunity. Thus, short- and long-term anti-DENV nAb titers in the sera of immunized mice were measured using a plaque reduction neutralizing test, as reported previously (Wang et al., 2019). The viral strains were DENV1 strain Hawaii, DENV2 strain Tr1751, DENV3 strain H87, and DENV4 strain H241, respectively. Over the short-term, sera from TetraME-vaccinated mice displayed broad neutralizing potency against heterologous DENVs. The nAb titers against DENV1–4 were 1: 160, 1: 226, 1: 174, and 1: 174, respectively, which were greater than their corresponding controls (Figure 1D, $P<0.01$). Similar to the dynamic changes in IgG titers, at 30 weeks after the last immunization, the nAb titers in the sera of the TetraME groups were 1: 104, 1: 135, 1: 123, and 1: 113, respectively, which were lower than their short-term titers. However, an anti-DENV nAb titer of 1: 10 implies protection against challenge (Zhang et al., 2015). Thus, these data indicate that the three EP doses of TetraME induced potent and long-term antibody responses with strong neutralizing activity.

Notwithstanding, the inevitable waning of nAb with time following TetraME vaccination raises concerns about whether the protective efficacy of the vaccine can persist long enough. Therefore, we performed DENV challenge experiments at week 30 after the last vaccination. Mice were challenged intracerebrally with either DENV1 (strain Hawaii) at a dose of 1×10^6 plaque-forming units (PFU), DENV2 (strain Tr 1751) at a dose of 200 PFU, DENV3 (strain H87) at a dose of 1×10^6 PFU, or DENV4 (strain H241) at a dose of 1×10^5 PFU. Each experiment was independently repeated three times. All TetraME-vaccinated mice showed only slight body weight losses, ranging from 3.4%–6.3% (Figure 1E–H, $P<0.01$) and survived the challenge with different DENV serotypes. In contrast, mice in the pV groups showed obvious body weight losses, ranging from 17.4%–19.2% after DENV challenge and all died, except for DENV3 group mice, which showed a 25% survival (2/8). This indicated that vaccination of mice with TetraME was sufficient to induce prolonged protective immunity against lethal challenge.

Over the last decade, progress with DNA vaccines has lagged behind that of others due to their limited immunogenicity (Kudlacek & Metz, 2019). To address these concerns, EP combined with multiple immunization strategies has been applied, resulting in dramatically improved immunogenicity and sustained expression of antigen-encoding DNA vaccines (Li & Petrovsky, 2016; Sheng et al., 2016). Here, our dengue DNA vaccine candidate (TetraME) generated robust DENV-specific cellular and humoral immune responses as well as long-term protective efficacy against four DENV serotypes in mice, which is beneficial in several aspects: firstly, DNA vaccination by intramuscular EP can generate significant antibody responses that persist for at least half a year (Babiuk et al., 2007) as well as antigen-specific T cell responses that persist for 40–60 weeks (Davis et al., 1995; Gurunathan et al., 1998); secondly, cross-protection among serotypes can be induced by monovalent

vaccines. For example, infection or immunization with one DENV serotype can confer substantial cross-protection against heterologous serotypes for an average duration of two years (Reich et al., 2013), consistent with our results. However, the characteristics of cross-immunity for each monovalent DNA component to other serotypes of DENV warrant further in-depth investigation. In the current study, we used immunocompetent BALB/c mice vaccinated with four monovalent prM/E-based DNA vaccine candidates (TetraME) and verified their long-term tetravalent protection up to 30-weeks post-vaccination. These data should provide a basis for further development and testing of this vaccine formulation in larger animals in combination with EP delivery.

In conclusion, we determined the duration and protection of resulting tetravalent antibodies by vaccinating mice with a tetravalent dengue DNA vaccine. For the first time, we demonstrated that DENV-specific nAb titers remained relatively constant and conferred full protection for up to 30 weeks after immunization. Thus, this study provides promising data for the further development of tetravalent DNA vaccines against dengue.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

R. W. designed and performed the experiments, analysed the data, and wrote the manuscript; F. J. Y. analyzed the data and reviewed the manuscript; X. Y. Z. and X. Z. L. helped with the experiments; D. Y. F. contributed reagents and materials; H. C. designed the research; J. A. principally designed the experiments, revised the manuscript, and directed the project. All authors read and approved the final version of the manuscript.

Ran Wang[1,2,#], Fu-Jia Yang[2,#], Xiao-Yan Zheng[3], Xian-Zheng Liao[2], Dong-Ying Fan[2], Hui Chen[2,*], Jing An[2,4,*]

[1] *Key Laboratory of Major Diseases in Children, Ministry of Education, National Clinical Research Center for Respiratory Diseases, Beijing Key Laboratory of Pediatric Respiratory Infection Diseases, Virology Laboratory, Beijing Pediatric Research Institute, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Research Unit of Critical Infection in Children, Chinese Academy of Medical Sciences, 2019RU016, Beijing 100045, China*
[2] *Department of Microbiology, School of Basic Medical Sciences, Chinese Capital Medical University, Beijing 100069, China*
[3] *Beijing Tropical Medicine Research Institute, Beijing Friendship Hospital, Second Clinical Medical College of Capital Medical University, Beijing 100050, China*
[4] *Center of Epilepsy, Beijing Institute for Brain Disorders, Beijing 100069, China*
[#]Authors contributed equally to this work
*Corresponding authors, E-mail: chenhuicxh@ccmu.edu.cn; anjing@ccmu.edu.cn

REFERENCES

Babiuk S, Tsang C, van Drunen Littel-van den Hurk S, Babiuk LA, Griebel PJ. 2007. A single hbsag DNA vaccination in combination with electroporation elicits long-term antibody responses in sheep. *Bioelectrochemistry*, **70**(2): 269−274.

Chen H, Zheng X, Wang R, Gao N, Sheng Z, Fan D, Feng K, Liao X, An J. 2016. Immunization with electroporation enhances the protective effect of a DNA vaccine candidate expressing prme antigen against dengue virus serotype 2 infection. *Clinical Immunology*, **171**: 41−49.

Davis HL, Schirmbeck R, Reimann J, Whalen RG. 1995. DNA-mediated immunization in mice induces a potent mhc class i-restricted cytotoxic t lymphocyte response to the hepatitis b envelope protein. *Human Gene Therapy*, **6**(11): 1447−1456.

Fatima K, Syed NI. 2018. Dengvaxia controversy: Impact on vaccine hesitancy. *Journal of Global Health*, **8**(2): 010312.

Fynan EF, Lu S, Robinson HL. 2018. One group's historical reflections on DNA vaccine development. *Human Gene Therapy*, **29**(9): 966−970.

Gurunathan S, Prussin C, Sacks DL, Seder RA. 1998. Vaccine requirements for sustained cellular immunity to an intracellular parasitic infection. *Nature Medicine*, **4**(12): 1409−1415.

Kudlacek ST, Metz SW. 2019. Focused dengue vaccine development: outwitting nature's design. *Pathogens and Disease*, **77**(1): ftz003.

Lee JS, Lourenco J, Gupta S, Farlow A. 2018. A multi-country study of dengue vaccination strategies with dengvaxia and a future vaccine candidate in three dengue-endemic countries: vietnam, thailand, and colombia. *Vaccine*, **36**(17): 2346−2355.

Li L, Petrovsky N. 2016. Molecular mechanisms for enhanced DNA vaccine immunogenicity. *Expert Review Vaccines*, **15**(3): 313−329.

Prompetchara E, Ketloy C, Thomas SJ, Ruxrungtham K. 2019. Dengue vaccine: global development update. *Asian Pacific Journal of Allergy and Immunology*.

Reich NG, Shrestha S, King AA, Rohani P, Lessler J, Kalayanarooj S, Yoon IK, Gibbons RV, Burke DS, Cummings DA. 2013. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. *Journal of the Royal Society Interface*, **10**(86): 20130414.

Sheng Z, Chen H, Feng K, Gao N, Wang R, Wang P, Fan D, An J. 2019. Electroporation-mediated immunization of a candidate DNA vaccine expressing dengue virus serotype 4 prm-e antigen confers long-term protection in mice. *Virologica Sinica*, **34**(1): 88−96.

Sheng Z, Gao N, Cui X, Fan D, Chen H, Wu N, Wei J, An J. 2016. Electroporation enhances protective immune response of a DNA vaccine against japanese encephalitis in mice and pigs. *Vaccine*, **34**(47): 5751−5757.

Shrivastava A, Tripathi NK, Dash PK, Parida M. 2017. Working towards dengue as a vaccine-preventable disease: challenges and opportunities. *Expert Opinion Biological Therapy*, **17**(10): 1193−1199.

Wang R, Liao X, Fan D, Wang L, Song J, Feng K, Li M, Wang P, Chen H, An J. 2018. Maternal immunization with a DNA vaccine candidate elicits specific passive protection against post-natal zika virus infection in immunocompetent balb/c mice. *Vaccine*, **36**(24): 3522−3532.

Wang R, Zheng X, Sun J, Feng K, Gao N, Fan D, Chen H, Jin X, An J. 2019. Vaccination with a single consensus envelope protein ectodomain sequence administered in a heterologous regimen induces tetravalent immune responses and protection against dengue viruses in mice. *Frontiers in Microbiology*, **10**: 1113.

Wilder-Smith A, Ooi EE, Horstick O, Wills B. 2019. Dengue. *The Lancet*, **393**: 350−363.

Zhang ZS, Weng YW, Huang HL, Zhang JM, Yan YS. 2015. Neutralizing antibodies respond to a bivalent dengue DNA vaccine or/and a recombinant bivalent antigen. *Molecular Medicine Reports*, **11**(2): 1009−1016.

Zheng X, Chen H, Wang R, Fan D, Feng K, Gao N, An J. 2017. Effective protection induced by a monovalent DNA vaccine against dengue virus (dv) serotype 1 and a bivalent DNA vaccine against dv1 and dv2 in mice. *Fronties in Cellular and Infection Microbiology*, **7**: 175.

CaptureProbe: a java tool for designing probes for capture Hi-C applications

DEAR EDITOR,

Many functional elements associated with traits and diseases are located in non-coding regions and act on distant target genes via chromatin looping and folding, making it difficult for scientists to reveal the genetic regulatory mechanisms. Capture Hi-C is a newly developed chromosome conformation capture technology based on hybridization capture between probes and target genomic regions. It can identify interactions among target loci and all other loci in a genome with low cost and high resolution. Here, we developed CaptureProbe, a user-friendly, graphical Java tool for the design of capture probes across a range of target sites or regions. Numerous parameters helped to achieve and optimize the designed probes. Design testing of CaptureProbe showed high efficiency in the design success ratio of target loci and probe specificity. Hence, this program will help scientists conduct genome spatial interaction research. CaptureProbe and source code are available at https://sourceforge.net/projects/captureprobe/.

Genome level studies on traits and diseases in different organisms have revealed that the majority of associated genetic loci are located in non-coding regions and are enriched in different regulatory signals, thus suggesting their regulatory functions (Maurano et al., 2012; Welter et al., 2014; Zhang et al., 2014). Regulatory elements can act on multiple genes and distant target genes via chromatin looping (Maston et al., 2006). Therefore, elucidation of the regulatory mechanisms of these non-coding loci is not reliable when applying simple assignment to the nearest genes. Chromosome conformation capture with high-throughput sequencing (Hi-C) allows for the identification of physical chromatin interactions across an entire genome (Lieberman-Aiden et al., 2009). However, the enormous complexity of Hi-C libraries makes it costly to obtain sufficient spatial resolution to detect interactions among specific elements. To circumvent these issues, capture Hi-C technology with capture probes was developed to reduce the target regions for sequencing in order to identify interactions among target loci and all other loci in a genome at low cost (Mifsud et al., 2015; Sahlén et al., 2015; Schoenfelder et al., 2015). This technology has been used extensively in different studies to reveal the regulatory mechanisms of traits or disease-associated loci in non-coding regions (Baxter et al., 2018; Mishra & Hawkins, 2017). The design of capture probes is a necessary prerequisite for capture Hi-C experiments and can be complex work for researchers without programming experience.

Several software tools have been designed for capture Hi-C probes, including CapSequm (Davies et al., 2016), HiCapTools (Anil et al., 2018), and GOPHER (Hansen et al., 2019). These toolkits are important in capture Hi-C-related analysis but cannot meet all requirements of diverse experiments. For instance, CapSequm, which is a web application for designing capture probes, can only process 1 000 positions at a time and provides very limited design parameters (e.g., probe length, restriction enzyme). HiCapTools was designed to find probes for target sites, but not for target regions, which are very common candidates for genetic research. In addition, HiCapTools contains limited parameters, which reduces its flexibility when considering specific DNA sequencing contexts. Furthermore, it is a command-line program and requires a series of input files, and thus is not very user friendly. The recently developed program GOPHER can design capture probes for both target sites and regions and includes a user-friendly graphic user interface (GUI). However, its capture probe design capacity is currently limited to human and mouse.

In this study, we developed CaptureProbe, a Java tool with a graphical user-friendly interface that can design capture probes for both target genetic sites and regions without species limitation. CaptureProbe is easy to use, only requires

simple input files, and provides abundant parameters for probe design. Moreover, it can also give detailed statistical information about design results. Comparisons between CaptureProbe and other existing tools showed that it provides rich software functions and shows better or equivalent performance in designing capture probes.

To achieve good performance in capturing informative ligation fragments, CaptureProbe designs probes based on the structural features of the Hi-C library. The Hi-C library consists of ligated restriction fragments originally in close spatial proximity in the nucleus (Lieberman-Aiden et al., 2009). Usually, these ligation fragments are sheared to a specific size range to ensure suitability for high-throughput sequencing. Therefore, CaptureProbe designs probes to capture both ends of the target restriction fragment (overlapping target sites or regions) and selects probes nearest to the end of the target restriction fragment. The program initially starts probe design from both ends of the target restriction fragment and moves inward by 1 bp for each cycle. To improve capture efficiency and specificity of probes, CaptureProbe calculates the GC content and missing and repeated bases (missing bases: n/N, repeated bases: marked in lowercase) in the probe sequence and chooses the first probe that meets all parameter limitations provided by the user. CaptureProbe can avoid redundancy probe sequences caused by target site overlap in the same restriction fragment.

Running CaptureProbe is very simple, requiring only the coordinate file of the target sites/regions and the sequence file (fasta format). CaptureProbe can limit any repeated sequences marked in the sequence files when designing the probes. Users can employ windows to directly specify the path of the required files and to set parameters. All real-time

configuration information can be printed for users to check progress. After running, CaptureProbe can print detailed information on the results of the capture probe design for users to evaluate the results. CaptureProbe will generate a series of result files for users to customize probes and to check the design state of each target site or region.

We systematically compared software function and probe design performance between CaptureProbe and other existing tools (Tables 1–2). As CapSequm function is limited, comparison analysis was not included. Both CaptureProbe and GOPHER showed rich functions and user-friendly GUI (Table 1).

Table 1 Functional comparisons among CaptureProbe and other tools

Tool	CaptureP-robe	HiCap-Tools	GOPHER
Supported species	All	All	Human, Mouse
Type of target loci	Site/Region	Site	Site/Re-gion
GUI	√	×	√
Detailed probe information	√	×	√
Probe GC content limitation	√	×	√
Probe missing base limitation	√	×	×
Repeated sequence limitation	√	√	×
Design margin limitation	√	×	√
Fragment size limitation	√	×	√
Fragment missing base limitation	√	×	×
Mapping score limitation	×	√	√

Table 2 Comparisons of capture probe design among CaptureProbe and other tools

Tool	CaptureProbe	HiCapTools	GOPHER
Testing site number (*n*)	20 000	20 000	20 000
Both ends with probes (%)	42.40	21.26	85.76
Only upstream with probe (%)	18.77	23.14	2.67
Only downstream with probe (%)	19.28	23.77	3.02
Total sites with probes (%)	80.45	68.17	91.45
Total sites with no probes (%)	19.57	31.83	8.56
Probe GC content <25% (%)	0.00	3.03	0.00
Probe GC content >65% (%)	0.00	0.34	0.00
Probe with extreme GC content (%)	0.00	3.37	0.00
Probe with unique alignment (%)	92.84	77.44	83.34
Probe with multiple aligments (%)	7.10	22.31	16.41
Probe with no alignment (%)	0.06	0.25	0.25

In this study, we only evaluated design performance for target sites as the mechanism is the same for target sites and regions. Twenty thousand random target sites (not from gap regions) in the human genome (hg38) were generated for testing. The same parameters were set for all tools: i.e., probe length, 120 bp; repeat sequence length, 6 bp; restriction

enzyme, Hand III; minimal fragment length, 300 bp; design margin size, 500 bp; probe GC content, 25%–65%; with all other parameters set using default values. Firstly, we compared the design success ratio among the three programs. GOPHER showed the highest design success ratio (91.45%), followed by CaptureProbe (80.45%), and finally

HiCapTools (68.17%). We next accessed the specificity of the probes, using BLASTN (Altschul et al., 1990) to align all probes to the genome sequence. CaptureProbe demonstrated the highest ratio of unique alignment (92.84%) among the programs (GOPHER: 83.34%, HiCapTools: 77.44%). As HiCapTools could not filter GC content in the probe sequences, partial probes of HiCapTools (3.37%) showed extremely high GC content (<25% or >65%), which did not match the efficient capture range (Agilent Technologies). Furthermore, we also found that small probes from GOPHER contained ambiguous characters (N).

Here, we present a very simple and user-friendly Java tool (CaptureProbe) that facilitates rapid capture probe design for target chromosome capture applications with no species limitation. CaptureProbe provides rich software functions and shows good probe design performance. Comparisons with existing software demonstrated that CaptureProbe has a good design success ratio and better probe specificity. CaptureProbe will be useful for a wide range of scientists studying genome spatial interactions.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.F.M., Y.P.Z., and H.B.X. designed the research. Y.F.M. implemented the Java code and analyzed the data. Y. F. M., Y. P. Z., and H. B. X. wrote the paper. A. C. A. and Y. B. S. revised and edited the manuscript. All authors read and approved the final version of the manuscript.

Yun-Fei Ma[1,2,3], Adeniyi C. Adeola[1], Yan-Bo Sun[1],
Hai-Bing Xie[1,*], Ya-Ping Zhang[1,4,*]

[1] *State Key Laboratory of Genetic Resources and Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*
[2] *Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China*
[3] *University of Chinese Academy of Sciences, Beijing 100049, China*
[4] *State Key Laboratory for Conservation and Utilization of Bio-resource, and Key Laboratory for Animal Genetic Diversity and Evolution of High Education in Yunnan Province, Yunnan University, Kunming, Yunnan 650091, China*
*Corresponding authors, E-mail: xiehb@mail.kiz.ac.cn; zhangyp@mail.kiz.ac.cn

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, **215**(3): 403−410.

Anil A, Spalinskas R, Åkerborg Ö, Sahlén P. 2018. HiCapTools: a software suite for probe design and proximity detection for targeted chromosome conformation capture applications. *Bioinformatics*, **34**(4): 675−677.

Baxter JS, Leavy OC, Dryden NH, Maguire S, Johnson N, Fedele V, Simigdala N, Martin LA, Andrews S, Wingett SW, Assiotis I, Fenwick K, Chauhan R, Rust AG, Orr N, Dudbridge F, Haider S, Fletcher O. 2018. Capture Hi-C identifies putative target genes at 33 breast cancer risk loci. *Nature Communications*, **9**(1): 1028.

Davies JO, Telenius JM, Mcgowan SJ, Roberts NA, Taylor S, Higgs DR, Hughes JR. 2016. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nature Methods*, **13**(1): 74−80.

Hansen P, Ali S, Blau H, Danis D, Hecht J, Kornak U, Lupiáñez DG, Mundlos S, Steinhaus R, Robinson PN. 2019. GOPHER: Generator of probes for capture Hi-C experiments at high resolution. *BMC Genomics*, **20**(1): 40.

Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**(5950): 289−293.

Maston GA, Evans SK, Green MR. 2006. Transcriptional regulatory elements in the human genome. *Annual Review of Genomics and Human Genetics*, **7**: 29−59.

Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA. 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science*, **337**(6099): 1190−1195.

Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA, Herman B, Happe S, Higgs A, Leproust E, Follows GA, Fraser P, Luscombe NM, Osborne CS. 2015. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics*, **47**(6): 598−606.

Mishra A, Hawkins RD. 2017. Three-dimensional genome architecture and emerging technologies: looping in disease. *Genome Medicine*, **9**(1): 87.

Sahlén P, Abdullayev I, Ramsköld D, Matskova L, Rilakovic N, Lötstedt B, Albert TJ, Lundeberg J, Sandberg R. 2015. Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biology*, **16**(1): 156.

Schoenfelder S, Furlan-Magaril M, Mifsud B, Tavares-Cadete F, Sugar R, Javierre BM, Nagano T, Katsman Y, Sakthidevi M, Wingett SW, Dimitrova E, Dimond A, Edelman LB, Elderkin S, Tabbada K, Darbo E, Andrews S, Herman B, Higgs A, Leproust E, Osborne CS, Mitchell JA, Luscombe NM, Fraser P. 2015. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Research*, **25**(4): 582−597.

Welter D, Macarthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, **42**(D1): D1001−1006.

Zhang X, Bailey SD, Lupien M. 2014. Laying a solid foundation for Manhattan--'setting the functional basis for the post-GWAS era'. *Trends in Genetics*, **30**(4): 140−149.